

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 11:25:54 ; Search time 13131 Seconds
(without alignments)
11406.220 Million cell updates/sec

Title: US-09-983-000a-1
Perfect score: 3091
Sequence: 1 cacacatacgcacgcacgat.....ttatagtagtaagtgtatataca 3091

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_scs.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2018.4	65.3	70276	9	AC073095	AC073095 Homo sapi
2	2008.8	65.0	159994	2	AC073471	AC073471 Homo sapi
3	1263.2	40.9	7941	6	CQ776705	CQ776705 Sequence
4	1263.2	40.9	7941	6	AR231695	AR231695 Sequence
5	1263.2	40.9	7941	6	AR231695	AR231695 Sequence
6	1263.2	40.9	7941	9	HUMPTPRZ	M93426 Human prote
7	1263.2	40.9	7941	11	G20044	G20044 sWSS1987 Er
8	1242.2	40.2	3350	9	HSU88967	U88967 Human tyros
9	1116.2	36.1	6807	6	CQ725791	I35776 Sequence 2
10	1113	36.0	6924	6	I35776	CQ777575
11	935.2	30.3	6887	6	CQ777575	AJ133130 Mus muscu
12	935.2	30.3	6887	10	MMU133130	AX683146 Sequence
13	926.4	30.0	6801	6	AX683146	U04998 Rattus norv
14	926.4	30.0	6801	10	RNU04998	U09357 Rattus norv
15	926.4	30.0	7851	10	RNU09357	AJ428208 Mus muscu
16	911.4	29.5	2784	10	MMU428208	BC070850 Xenopus l
17	613.2	19.8	5281	5	BC070850	AB045238 Xenopus l
18	608.8	19.7	5284	5	AB045238	AB045237 Xenopus l
19	608.8	19.7	7725	5	AB045237	

20	401.6	13.0	1650	5	BC078205	BC078205 Danio rer
c 21	216.2	7.0	185859	2	AC133599	AC133599 Mus muscu
22	205.4	6.6	158090	9	AC006353	AC006353 Homo sapi
c 23	205.4	6.6	173354	2	AC146003	AC146003 Pan trogl
24	199.8	6.5	1605	4	BOVRPTPB	L29425 Bovine tyro
25	190.4	6.2	229597	2	AC094129	AC094129 Rattus no
26	189	6.1	4338	6	AR069915	AR069915 Sequence
27	189	6.1	4338	6	BD145125	BD145125 Novel rec
28	189	6.1	4338	6	I23305	I23305 Sequence 1
29	189	6.0	4620	10	HUMPTPRG	L09247 Human recep
30	185.4	6.0	5403	5	GGU38349	L09562 Mouse recep
31	185.2	5.8	2581	10	AY177706	U38349 Gallus gall
32	180.6	5.8	3721	10	AY177705	AY177706 Rattus no
33	180.6	5.8	3721	10	AY177705	AY177705 Rattus no
34	180.6	5.8	4391	10	AY177704	AY177704 Rattus no
35	180.6	5.8	4478	10	AY177703	AY177703 Rattus no
c 36	162.8	5.3	203503	2	AC116273	AC116273 Rattus no
37	161.8	5.2	960	10	S57181	S57181 Ptpg=recep
38	156.4	5.1	268	9	HS157F11R	Z57006 H.sapiens C
39	151	4.9	201	11	BV175013	BV175013 sqnm76079
40	123.6	4.0	5617	10	BC079595	BC079595 Mus muscu
41	103	3.3	1755	5	BC072918	BC072918 Xenopus l
c 42	101.8	3.3	46546	2	AC111161	AC111161 Homo sapi
43	90.2	2.9	3369	6	CQ721425	CQ721425 Sequence
c 44	85.2	2.8	237318	2	CR752641	CR752641 Danio rer
45	83.8	2.7	178398	2	CR388022	CR388022 Danio rer

ALIGNMENTS

RESULT 1	AC073095	Homo sapiens BAC clone RP11-384A20 from 7, complete sequence.	DNA	linear	PRI 08-OCT-2003
LOCUS	AC073095	HTG	70276 bp		
DEFINITION	AC073095	HTG			
ACCESSION	AC073095.4	GI:21428060			
VERSION	AC073095.4	GI:21428060			
KEYWORDS	HTG				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	Sulston,J.E. and Wilson,R.			
AUTHORS	1	(bases 1 to 70276)			
TITLE	1	Toward a complete human genome sequence			
JOURNAL	1	Genome Res. 8 (11), 1097-1108 (1998)			
MEDLINE	1	99063792			
PUBMED	1	9847074			
REFERENCE	2	(bases 1 to 70276)			
AUTHORS	2	Cotton,M., Glaser,E. and Pearman,C.			
TITLE	2	The sequence of Homo sapiens BAC clone RP11-384A20			
JOURNAL	2	Unpublished (2001)			
REFERENCE	3	(bases 1 to 70276)			
AUTHORS	3	Waterston,R.H.			
TITLE	3	Direct Submission			
JOURNAL	3	Submitted (08-JUN-2000) Genome Sequencing Center, Washington			
REFERENCE	4	(bases 1 to 70276)			
AUTHORS	4	Waterston,R.H.			
TITLE	4	Direct Submission			
JOURNAL	4	Submitted (13-DEC-2000) Genome Sequencing Center, Washington			
REFERENCE	5	(bases 1 to 70276)			
AUTHORS	5	Waterston,R.			
TITLE	5	Direct Submission			
JOURNAL	5	Submitted (09-MAY-2001) Department of Genetics, Washington			
REFERENCE	6	(bases 1 to 70276)			
AUTHORS	6	Waterston,R.H.			
TITLE	6	Direct Submission			
JOURNAL	6	Submitted (15-JUN-2002) Genome Sequencing Center, Washington			

University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
7 (bases 1 to 70276)
Waterston.R.
Direct Submission
Submitted (29-APR-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
8 (bases 1 to 70276)

COMMENT

NOTICE: This sequence may not represent the entire insert of this clone! It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality 30); an attempt was made to resolve all sequencing problems, as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director) John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:sqeen@nhgri.nih.gov>, or see <http://genome.wustl.edu>

SOURCE INFORMATION:

The R6c11 human BAC library was made from the blood of one male donor as described by Oregawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tatchell, M., Cataneese, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP5-1062J16, 200 bp overlap the clone sequenced to the right is RP5-1049N15, 200 bp overlap. Actual start of this clone is at base position 101340 of RP5-1062J16 actual end is at base position 34395 of RP5-1049N15.

FEATURES

source

```
/organism="Homo sapiens"  
/mol_type="genomic DNA"  
/db_xref="taxon:9606"  
/chromosome="7"  
/clone="RP11-384A20"  
/clone_lib="RPC1-11"  
678. 972  
/rpt_family="Alu"  
1170. 1279  
/rpt_family="L2"
```

repeat region

repeat region

repeat_region	2555..3252	/rpt_family="L1"
repeat_region	3278..3300	/rpt_family="AT_rich"
repeat_region	4237..4361	/rpt_family="L2"
repeat_region	4365..4482	/rpt_family="Alu"
repeat_region	4487..4721	/rpt_family="L2"
repeat_region	4743..4936	/rpt_family="MIR"
repeat_region	4944..5311	/rpt_family="ERV1"
repeat_region	5315..5409	/rpt_family="ERV1"
repeat_region	5518..5671	/rpt_family="CR1"
repeat_region	5674..5731	/rpt_family="ERV1"
repeat_region	5736..6005	/rpt_family="CR1"
repeat_region	6006..6364	/rpt_family="MaLR"
repeat_region	6365..6404	/rpt_family="CR1"
repeat_region	7983..8286	/rpt_family="Alu"
misc_feature	8012..8285	/note="CpG_island (%GC=60.6, o/e=0.77, #CpGs=20)"
repeat_region	8343..8644	/rpt_family="Alu"
repeat_region	9957..10789	/rpt_family="L2"
repeat_region	11461..11495	/rpt_family="(TTTG)n"
repeat_region	13546..13640	/rpt_family="L2"
repeat_region	13679..13770	/rpt_family="MIR"
repeat_region	14013..14624	/rpt_family="L2"
repeat_region	14850..15153	/rpt_family="Alu"
repeat_region	16080..16372	/rpt_family="Alu"
repeat_region	18199..18513	/rpt_family="Alu"
repeat_region	18606..18894	/rpt_family="Alu"
repeat_region	19156..19389	/rpt_family="ERV1"
repeat_region	19364..19415	/rpt_family="L2"
repeat_region	19481..19864	/rpt_family="MaLR"
repeat_region	20205..20236	/rpt_family="ERV1"
repeat_region	20237..20613	/rpt_family="MaLR"
repeat_region	20614..20673	/rpt_family="L2"
repeat_region	20697..21120	/rpt_family="L2"
repeat_region	21121..21384	/rpt_family="MaLR"
repeat_region	21385..21412	/rpt_family="(TTTG)n"
repeat_region	21413..21789	/rpt_family="MaLR"
repeat_region	21871..22000	/rpt_family="MER1_type"
repeat_region	22018..22260	

[illegible]

QY 2992 GTAAATGTTGACCTGTTATTTTGGATAAATTAAGAGTTGGTTTAATTTGTTATTTATTTCC 3051
Db 54329 GTAAATGTTGACAGCTGTTATTTTGGATAAATTAAGAGTTGGTTTAATTTGTTATTTATTTCC 54388
QY 3052 TCTTTTCAGCCGCCAAAGCATTATGTAGTAAGTGATATACA 3091
Db 54389 TCTTTTCAGCCGCCAAAGCATTATGTAGTAAGTGATATACA 54428

RESULT 2
AC073471 159994 bp DNA linear HTG 25-OCT-2000
LOCUS Homo sapiens chromosome 7 clone RP11-593H15, WORKING DRAFT
DEFINITION SEQUENCE, 6 unordered pieces.
AC073471
VERSION AC073471.4 GI:10946569
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 159994)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 159994)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Oct 23, 2000 this sequence version replaced gi:9795985.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H_NH0593H15
----- Summary Statistics -----
Sequencing vector: M13; 70k
Chemistry: Dye-terminator ET; 70k of reads
Chemistry: Dye-terminator Big Dye; 0k of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 154286 bases at least Q40
Consensus quality: 156243 bases at least Q30
Consensus quality: 157367 bases at least Q20
Insert size: 150000; agarose-fp
Insert size: 159494; sum-of-contigs
Quality coverage: 5.07 in Q20 bases; agarose-fp
Quality coverage: 5.08 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 5146: contig of 5146 bp in length
* 5147 5246: gap of unknown length
* 5247 16991: contig of 11745 bp in length
* 16992 17091: gap of unknown length
* 17091 34473: contig of 17382 bp in length
* 34473 34573: gap of unknown length
* 34573 55892: contig of 21319 bp in length
* 55892 55992: gap of unknown length
* 55992 105455: contig of 49463 bp in length
* 105455 105555: gap of unknown length
* 105555 159994: contig of 54439 bp in length.

FEATURES		Location/Qualifiers	
source		1..159994	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		/chromosome="7"	
		/clone="RP11-593H15"	
misc_feature		1..5146	
		/note="assembly_name:Contig2"	
misc_feature		5247..16991	
		/note="assembly_name:Contig3"	
misc_feature		17092..34473	
		/note="assembly_name:Contig4"	
misc_feature		34574..55892	
		/note="assembly_name:Contig5"	
misc_feature		55993..105455	
		/note="assembly_name:Contig6"	
misc_feature		105556..159994	
		/note="assembly_name:Contig7"	
ORIGIN			
Query Match		65.0%; Score 2008.8; DB 2; Length 159994;	
Best Local Similarity		99.7%; Pred. No. 0;	
Matches 2013; Conservative		0; Mismatches 7; Indels 0; Gaps 0;	
QY	1072	GCAGTTTGTAGTTTCAGAACCCAGAAATGTTTCAGGCTGACCCAGAGAAATTATACGAGCCTT	1131
Db	77763	GAAGTTTGTAGTTTCAGAACCCAGAAATGTTTCAGGCTGACCCAGAGAAATTATACGAGCCTT	77822
QY	1132	CTTGTTCATACGGAAGACCTCCGAGTCGTTTATGATACCATGATGAGAAAGTTTGCAAGTT	1191
Db	77823	CTTGTTCATACGGAAGACCTCCGAGTCGTTTATGATACCATGATGAGAAAGTTTGCAAGTT	77882
QY	1192	TTGTACACGAGTTGGATGGAGAGGCCAAACCAAGCATGAATTTTGCAGAGATGGCTAT	1251
Db	77883	TTGTACACGAGTTGGATGGAGAGGCCAAACCAAGCATGAATTTTGCAGAGATGGCTAT	77942
QY	1252	CAAGACTTGGTAACTATATCATGATCAGTTGTTTTCATAGGGTAAACATTATTAATTTTC	1311
Db	77943	CAAGACTTGGTAACTATATCATGATCAGTTGTTTTCATAGGGTAAACATTATTAATTTTC	78002
QY	1312	CAAGGTAAGAACTTACAAATGGTTGTATATTTTCTCCATCTACCTTTTAGACTTTATG	1371
Db	78003	CAAGGTAAGAACTTACAAATGGTTGTATATTTTCTCCATCTACCTTTTAGACTTTATG	78062
QY	1372	TGAAGTGGGGTAGGCTGAGTATTTTAAATTTTAAAAAATTTTAAATTTAGAGCTAT	1431
Db	78063	TGAAGTGGGGTAGGCTGAGTATTTTAAATTTTAAAAAATTTTAAATTTAGAGCTAT	78122
QY	1432	ACTAAATTTATGTTTAAAGTTACATTTAATTTAAATTTGATATCATAACTTTGCCAACAATA	1491
Db	78123	ACTAAATTTATGTTTAAAGTTACATTTAATTTAAATTTGATATCATAACTTTGCCAACAATA	78182
QY	1492	CACATATAGAGTAGATACATATGACTTATGAACCTGGAGATCATTTAGTGTGGCCCTTCTTA	1551
Db	78183	CACATATAGAGTAGATACATATGACTTATGAACCTGGAGATCATTTAGTGTGGCCCTTCTTA	78242
QY	1552	AGATTCAGTTGTAGAAATAGTGCAGAACTCTCAGTGCCTCGATACATTTTATTTGTGTC	1611
Db	78243	AGATTCAGTTGTAGAAATAGTGCAGAACTCTCAGTGCCTCGATACATTTTATTTGTGTC	78302
QY	1612	TTCCATTTACGCTATATCCAGCACAGGAAAGTAGAGTAGGGGACATACAAGTCTCTTTTGT	1671
Db	78303	TTCCATTTACGCTATATCCAGCACAGGAAAGTAGAGTAGGGGACATACAAGTCTCTTTTGT	78362
QY	1672	TGCACCAAAAAATTTTCAGATAAACAGCTGGGAAGTCAATGATGGGTGAGAACTTTTGGGA	1731
Db	78363	TGCACCAAAAAATTTTCAGATAAACAGCTGGGAAGTCAATGATGGGTGAGAACTTTTGGGA	78422
QY	1732	TGTAAGAAACATTTCTTACAAAAAGATCCACCCCTGCCTCCCTCCACGAGCGATCGGA	1791
Db	78423	TGTAAGAAACATTTCTTACAAAAAGATCCACCCCTGCCTCCCTCCACGAGCGATCGGA	78482


```
Db 901 |||||ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTT|||TGTGAAGTTCTTACAAATGCAACA 960
Qy 961 TCTGGTTATGTCATGCTGATGACTTACAAAACAATTTTCGAGAGCAACAGTACAA 1020
Db 961 TCTGGTTATGTCATGCTGATGACTTACAAAACAATTTTCGAGAGCAACAGTACAA 1020
Qy 1021 TTCTCTAGACAGGTGTTTCTCTATACATCTGAAAGGAGAGATTTCATGAAGCAGTTTGT 1080
Db 1021 TTCTCTAGACAGGTGTTTCTCTATACATCTGAAAGGAGAGATTTCATGAAGCAGTTTGT 1080
Qy 1081 AGTTCCAGAACGAGAAATGTTTTCAGGCTGACCCAGAGAAATATACAGGCTTCTTGTACA 1140
Db 1081 AGTTCCAGAACGAGAAATGTTTTCAGGCTGACCCAGAGAAATATACAGGCTTCTTGTACA 1140
Qy 1141 TGGGAAGACCTCGAGTGGTTTATCATACATGTAAGAGGAGTTTGAGAGTTTGTACCA 1200
Db 1141 TGGGAAGACCTCGAGTGGTTTATCATACATGTAAGAGGAGTTTGAGAGTTTGTACCA 1200
Qy 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGAAATTTTTCAGAGTGGCTATCAAGACTTG 1260
Db 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGAAATTTTTCAGAGTGGCTATCAAGACTTG 1260
Qy 1261 GTAACAT 1268
Db 1261 GGTGCTAT 1268

RESULT 5
AX334330 7941 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 4839 from Patent WO0194629.
DEFINITION AX334330
ACCESSION AX334330
VERSION AX334330.1 GI:18125049
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrikan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 4839 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source
1..7941
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 40.9%; Score 1263.2; DB 6; Length 7941;
Best Local Similarity 99.8%; Pred. No. 7.4e-230;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CACATATGCGACGACGATCTCCTTGGCTCCCTCCCTCTCCACTCTGAGAGGAGGAGCGCA 60
Db 1 CACATATGCGACGACGATCTCCTTGGCTCCCTCCCTCTCCACTCTGAGAGGAGGAGCGCA 60
Qy 61 CAAAAAACAATTTCTTGGCTCCCTCCCTCCCTCTCCACTCTGAGAGGAGGAGCGCA 120
Db 61 CAAAAAACAATTTCTTGGCTCCCTCCCTCCCTCTCCACTCTGAGAGGAGGAGCGCA 120
Qy 121 CGGCGAGGGCGCGAGACCGTCTGAAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180
Db 121 CGGCGAGGGCGCGAGACCGTCTGAAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180
Qy 181 CAGCTCCTCTGTGTTTGGCGCTGAGTTGGGCTAATGATATCTACAGACCAAGAGAAA 240
Db 181 CAGCTCCTCTGTGTTTGGCGCTGAGTTGGGCTAATGATATCTACAGACCAAGAGAAA 240
```

```
Qy 241 CTTGTTGAGAGATTTGGCTGCTCTATACAGGAGCACTGAATCAAAAAAATTTGGGGAAG 300
Db 241 CTTGTTGAGAGATTTGGCTGCTCTATACAGGAGCACTGAATCAAAAAAATTTGGGGAAG 300
Qy 301 AATATATCAACATGTAATAGCCAAAACAATCTCTATCAATATTTGATGAAGATCTTACA 360
Db 301 AATATATCAACATGTAATAGCCAAAACAATCTCTATCAATATTTGATGAAGATCTTACA 360
Qy 361 CAAGTAATGTGAATCTTAAAGAACTTTAAATTTTCAGAGGTTGGGATAAAACATCAATGGA 420
Db 361 CAAGTAATGTGAATCTTAAAGAACTTTAAATTTTCAGAGGTTGGGATAAAACATCAATGGA 420
Qy 421 AACACATTTCAATCAATACCTGGGAAAACAGTGGAAATTAATCTCACTAAATGACTACCGT 480
Db 421 AACACATTTCAATCAATACCTGGGAAAACAGTGGAAATTAATCTCACTAAATGACTACCGT 480
Qy 481 GTACGCGAGGAGTTTCAGAAATGGTGTAAAGCAAGCAAGATAAATCTTTCACCTGGGGA 540
Db 481 GTACGCGAGGAGTTTCAGAAATGGTGTAAAGCAAGCAAGATAAATCTTTCACCTGGGGA 540
Qy 541 AATGCAATATGTCATCTGATGATCAGAGCATAGTTTAGAAGGACAAAATTTTCCACTT 600
Db 541 AATGCAATATGTCATCTGATGATCAGAGCATAGTTTAGAAGGACAAAATTTTCCACTT 600
Qy 601 GAGATGCAAAATCTACTGCTTTGATGCGGACCGATTTTCAAGTTTTGGAGGAGCAGTCAAA 660
Db 601 GAGATGCAAAATCTACTGCTTTGATGCGGACCGATTTTCAAGTTTTGGAGGAGCAGTCAAA 660
Qy 661 GGAAGAGGAGTTAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAAATTTT 720
Db 661 GGAAGAGGAGTTAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAAATTTT 720
Qy 721 GATTTCAAAGCGATTTATTGATGGAGTGGAAAGTGTAGTCTGTTTGGAGGACGAGTCT 780
Db 721 GATTTCAAAGCGATTTATTGATGGAGTGGAAAGTGTAGTCTGTTTGGAGGACGAGTCT 780
Qy 781 TTAGATCCATTCATPACTGTTGAACCTTCTGCAAACTCAACTGCAAGATTTTACATTTAC 840
Db 781 TTAGATCCATTCATPACTGTTGAACCTTCTGCAAACTCAACTGCAAGATTTTACATTTAC 840
Qy 841 AATGGCTCATGATGATCTCTCTCCCTGACAGACACAGTTGATGATGATTTTAAAGAT 900
Db 841 AATGGCTCATGATGATCTCTCTCCCTGACAGACACAGTTGATGATGATTTTAAAGAT 900
Qy 901 ACAGTTAGCATCTCTGAAAGCAGTGGCTGTTTGTGAAGTTCTTACAAATGCAACA 960
Db 901 ACAGTTAGCATCTCTGAAAGCAGTGGCTGTTTGTGAAGTTCTTACAAATGCAACA 960
Qy 961 TCTGGTTATGTCATGCTGATGGACTACTTACAAAACAATTTTCGAGAGCAACAGTACAA 1020
Db 961 TCTGGTTATGTCATGCTGATGGACTACTTACAAAACAATTTTCGAGAGCAACAGTACAA 1020
Qy 1021 TTCTCTAGACAGGTGTTTCTCTATACATCTGAAAGGAGAGATTTCATGAAGCAGTTTGT 1080
Db 1021 TTCTCTAGACAGGTGTTTCTCTATACATCTGAAAGGAGAGATTTCATGAAGCAGTTTGT 1080
Qy 1081 AGTTCCAGAACGAGAAATGTTTTCAGGCTGACCCAGAGAAATATACAGGCTTCTTGTACA 1140
Db 1081 AGTTCCAGAACGAGAAATGTTTTCAGGCTGACCCAGAGAAATATACAGGCTTCTTGTACA 1140
Qy 1141 TGGGAAGACCTCGAGTGGTTTATGATACATGTAAGAGGAGTTTGAGAGTTTGTACCA 1200
Db 1141 TGGGAAGACCTCGAGTGGTTTATGATACATGTAAGAGGAGTTTGAGAGTTTGTACCA 1200
Qy 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGAAATTTTTCAGAGTGGCTATCAAGACTTG 1260
Db 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGAAATTTTTCAGAGTGGCTATCAAGACTTG 1260
Qy 1261 GTAACAT 1268
Db 1261 GGTGCTAT 1268
```

RESULT 6	
HUMPTPRZ	
LOCUS	7941 bp mRNA linear PRI 08-JAN-1995
DEFINITION	Human protein tyrosine phosphatase zeta-polypeptide (PTPRZ) mRNA, complete cds.
ACCESSION	M93426
VERSION	M93426.1 GI:190743
KEYWORDS	carbonic anhydrase-related transmembrane protein; protein tyrosine phosphatase zeta-polypeptide.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 7941)
TITLE	Krueger,N.X. and Saito,H. A human transmembrane protein-tyrosine-phosphatase, PTP zeta, is expressed in brain and has an N-terminal receptor domain homologous to carbonic anhydrases
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 89 (16), 7417-7421 (1992)
MEDLINE	92366472
PUBMED	1323835
COMMENT	Original source text: Homo sapiens (tissue library: lambda gt10 and gt11) fetus brain cDNA to mRNA.
FEATURES	Location/Qualifiers
source	1..7941
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/map="Unassigned"
	/tissue_type="brain"
	/dev_stage="fetus"
	/tissue_lib="lambda gt10 and gt11"
gene	1..7941
CDS	/gene="PTPRZ"
	148..7092
	/note="carbonic anhydrase-like domain"
	/codon_start=1
	/product="protein tyrosine phosphatase zeta-polypeptide"
	/protein_id="AAA60225.1"
	/db_xref="GI:190744"
	/db_xref="GDB:G00-127-353"
	/translation="WRILKRLFLICLLCVLRDWANGYRQRLVVEIGWSYTGAL NOKWKKYPTCNPKSPKPINIDELQVNVNKLKLFQGWKDTSLNFTLHNTGTV EINLTNDRVSGGVSEWVFASKITFHWGKNMSSDGEHSELEQKPLEMIQYCPDA DRFSFEAVKGRKRLALSILFEVGTENLDFKALIDGVESVRFRGQKALDPPILL NLLPNSDTKYIYNGSLTSPCTDTPVIMVFKDTSVISESOLAVFCEVLTMQSGYWM LMDYLQNNFRQYKFSRQVFSSTYKKEEIHAEVCSSEPNVQADPENYTSLLVTVR PRVVDYMIERFVLVYQDLGDDOTKHEFLTDGYQDLGAILNNLLPNMSVYVLIWAIC TNGLYKYSDDLIVDMPTDNPBLDLFPELIGTEERIKEEERKDIESGALVNPGRSA TNQIKKGPQLSTTHYNRIGTKNEAKTNRSPTRGSEFGKGVDPNTSLNSTQPTV KLATEKOISLTQVTELPHPHTVGTSTASLNDGSKTVLRSPPHMLSTGAEISLST EYEEESLTSFKLDTGADSSGSPATSAIPFI SENISQGYIFSENPEITTYDVLIP ESARNASDSTSSGSESLKDPMEGNWVFPSSDITAQDPVSGRESFLQNTYEIR VDESEKTTQPVAGVMSQGVTDLEMPHYSTPAYPTEVTPHAPTSPSRQDLVST VNVVYQSTQPVNGETPQPSYSEVFPPLVTPILLNQILNTPHAPTSPASSDSLHATP VFPVDVFEILSGIDGAPLPLPSSASFSSELFRHLHIVSQILPQVTSATESKVEL HASLPVAGGDLNLPQLAQSDVLTTHAASETFEGSESGVLYKTLMFQSVPEPSSD AMHARSGDPPFSLALSDNEGSHIFTVYSYSAIPVHDSVGTYQGLSPGSPHIP KSLIITPTPALQPTHALSGDEWSGASSDFLLPDTDGLTALNINSPVSAEFTVT TSVFGDDNKALSKEBIYGNETELQIPSFNEMVYPSSTVMNPNDVNVKNLASLOET SVSISSTKMPGSLAHTTKVFDHEISQVPENNFVQPTHTVSOQSDTSLKPLSA NSEPASSPASSEMSPQLLFVETSASFTEVLLQPSFOASDVDTLLKTLVLPVPS DPILVETPKVDKISSTMLHLIVSNSASENMLHSTSVFVDSPTSMHSHASLQGLTI SYASEKYPVLLKESHSQVPSLYSNDELFPQNTANLEINQAHPKGRHVFATPVLSID EPLANTINKLTHSDEILTSTKSVTVKVFAGIPTVASDTFVSTDSHPVIGNHVAITA VSPHRDGSCTVLLFPKSAISLSHSAKSADAGLVGGEGDGDTDGDDDDDDSDGL SIHKMCSGVSYREKQKWNDSHTHESLMDQNNPLISYLSLSENSENRVTSVSDSQ TGMDSGPKSPSANGLSQKNDGKEENDIQGSALLPLSPSKAHAVLTSDESGSQ GTSDSLNNETSTDFSFADTNEKADGILAAAGDSEITPFPQSPSTSVSSNSVFHV SEAEASNSHESRPIGLAEGLSEKKAIVPLIVSALTFICLVVLVGLIIVYRKCQFA HFYLEDSTSPRVITPPTPIPIISDVVGAIPIKHPKPHADLHASSGSEFEFTLKEF YQEVQSCVTDLGITADSSNHPDNKKNRYINIVAYDHSRVKLAQLAEKDGKLTIDYINA

```
Db 841 AATGGCTCATGTGACATCTCTCCCTGCACAGACACATGCTGATGCTGTTTAAAGAT 900
Qy 901 ACAGTTAGCATCTCTGAAGCCAGTTGGCTGTTTTTGTGAAGTTCTTACAATGCAACAA 960
Db 901 ACAGTTAGCATCTCTGAAGCCAGTTGGCTGTTTTTGTGAAGTTCTTACAATGCAACAA 960
Qy 961 TCTGGTTATGTCTATGCTGATGGAAGTACTTACAAAACAAATTTTCGAGAGCAACAGTACAG 1020
Db 961 TCTGGTTATGTCTATGCTGATGGAAGTACTTACAAAACAAATTTTCGAGAGCAACAGTACAG 1020
Qy 1021 TTCTCTAGACAGTGTGTTTCTCATACATCTGGAAGGAGAGATTCATGACGAGTTGT 1080
Db 1021 TTCTCTAGACAGTGTGTTTCTCATACATCTGGAAGGAGAGATTCATGACGAGTTGT 1080
Qy 1081 AGTTTCAGAACCCAGAAAATGTTTCAGGCTGACCCAGAGAAATATACCAAGCTTCTGTTACA 1140
Db 1081 AGTTTCAGAACCCAGAAAATGTTTCAGGCTGACCCAGAGAAATATACCAAGCTTCTGTTACA 1140
Qy 1141 TGGGAAGACCTCGAGTCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
Db 1141 TGGGAAGACCTCGAGTCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
Qy 1201 CAGTTGGATGGAGAGACCAACCAAGCATGAATTTTTCAGAGTGGCTATCAAGACTTG 1260
Db 1201 CAGTTGGATGGAGAGACCAACCAAGCATGAATTTTTCAGAGTGGCTATCAAGACTTG 1260
Qy 1261 GTAACTAT 1268
Db 1261 GGTGCTAT 1268
```

RESULT 7

```
G20044
LOCUS G20044
DEFINITION G20044 7941 bp DNA linear STS 28-SEP-1998
G20044.1 GI:1254743
STSS. Homo sapiens (human)
SOURCE Homo sapiens
```

```
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 7941)
Bouffard G.G., Iyer, L.M., Idol, J.R., Braden, V.V., Cunningham, A.P.,
Weintraub, J.A., Mohr-Tidwell, R.M., Peluso, D.C., Fulton, R.S.,
Leckie, M.P. and Green, E.D.
A collection of 1814 human chromosome 7-specific STSS
Genome Res. 7 (1), 59-64 (1997)
97189344
PUBMED 9037602
2 (Bases 1 to 7941)
Green, E.D.
Human chromosome 7 STSS (1997)
Unpublished (1997)
Synonyms: PTPRZ
GDB: GDB:3754378
GDB_DSEG: PTPRZ
Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
Tel: 3014020201
Fax: 3014024735
```

```
EMAIL: egreen@nhgri.nih.gov
PRIMER A: CTGGCTTACTCTACCCAG
PRIMER B: GAACAATGTTGGTAAGTG
STS size: 152
PCR Profile:
  Presoak: 0 degrees C for 0.00 minute(s)
  Denaturation: 92 degrees C for 0.17 minute(s)
  Annealing: 55 degrees C for 1.00 minute(s)
```

```
Polymerization: 72 degrees C for 1.00 minute(s)
PCR Cycles: 35
Thermal Cycler: PerkinElmer 9600
Protocol:
  Template: 30-100 ng
  Primer: each 1 uM
  dNTPs: each 200 uM
  Taq Polymerase: 0.05 units/uL
  Total Vol: 10 uL
Buffer:
  MgCl2: 2.5 mM
  KCl: 50 mM
  Tris-HCl: 10 mM
  pH: 8.3
```

This STS was developed from sequence determined by another investigator. See GenBank record: M93426 For additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgri.nih.gov/DIR/GB/CHR7>. Also see Genomics 11:548-64 (1991) [MUID=92128937].

FEATURES

```
source
1..7941
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="7"
/clone_lib="Eric D. Green"
1..7941
/gene="PTPRZ"
7705..7856
/gene="PTPRZ"
7705..7722
/gene="PTPRZ"
complement(7837..7856)
primer_bind
primer_bind
ORIGIN
```

```
Query Match 40.9%; Score 1263.2; DB 11; Length 7941;
Best Local Similarity 99.8%; Pred. No. 7,4e-230;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CACACATACGACGACGATCTCACTCGATCTATACACTCGAGGATTTAAACAAACAAA 60
Db 1 CACACATACGACGACGATCTCACTCGATCTATACACTCGAGGATTTAAACAAACAAA 60
Qy 61 CAAAAAACAATTTCTTCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 120
Db 61 CAAAAAACAATTTCTTCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 120
Qy 121 CGCGAGGGCGCGCAGACCGTCTGGAATGGAATCCTTAAAGCGTTTCTCGCTTGCATT 180
Db 121 CGCGAGGGCGCGCAGACCGTCTGGAATGGAATCCTTAAAGCGTTTCTCGCTTGCATT 180
Qy 181 CAGCTCCTCTGTGTTGTCGCTCGGATTTGGGCTAATGGATACAGACAGAGAGAAA 240
Db 181 CAGCTCCTCTGTGTTGTCGCTCGGATTTGGGCTAATGGATACAGACAGAGAGAAA 240
Qy 241 CTGTTGAAGAGATTTGGCTGCTCTATACAGGAGCATGAATCAAAAATTTGGGGAAG 300
Db 241 CTGTTGAAGAGATTTGGCTGCTCTATACAGGAGCATGAATCAAAAATTTGGGGAAG 300
Qy 301 AAATATCAACATGTAATAGCCCAAAACAATCTCTATCAATATTTGATGAAGATCTTACA 360
Db 301 AAATATCAACATGTAATAGCCCAAAACAATCTCTATCAATATTTGATGAAGATCTTACA 360
Qy 361 CAAGTAAATGTGAATCTTTAAGAACTTTAAATTTTCAGGGTTGGGATAAAACATCTTGA 420
Db 361 CAAGTAAATGTGAATCTTTAAGAACTTTAAATTTTCAGGGTTGGGATAAAACATCTTGA 420
Qy 421 AACACATTTCAATCAACATCTGGGAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
Db 421 AACACATTTCAATCAACATCTGGGAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
```


Db 548 ACAAGTAATGTAATCTTAAGAACTTAATTTACGGTGGGTAATAAATCATCTTGG 607
Qy 420 AAACACATTCATTCTAATCAACACTGGGAAACAGTGGAAATTAATCTCACTTAATGACTACCG 479
Db 608 AAACACATTCATTCTAATCAACACTGGGAAACAGTGGAAATTAATCTCACTTAATGACTACCG 667
Qy 480 TGTACAGCGAGGAGTTTCAGAAATGTTTAAAGCAAGCAAGATAAATCTTTTCACTGGGG 539
Db 668 TGTACAGCGAGGAGTTTCAGAAATGTTTAAAGCAAGCAAGATAAATCTTTTCACTGGGG 727
Qy 540 AAAATGCAATATGTCATCTGATGAGATCAGAGCATATGTTTAAAGCAAGCAAGATAAATCTTTTCACT 599
Db 728 AAAATGCAATATGTCATCTGATGAGATCAGAGCATATGTTTAAAGCAAGCAAGATAAATCTTTTCACT 787
Qy 600 TCAGATGCAAAATCTACTGCTTTGATGCGGACCGATTTTCAAGTTTGGAGAGCAGTCAA 659
Db 788 TGGATGCAAAATCTACTGCTTTGATGCGGACCGATTTTCAAGTTTGGAGAGCAGTCAA 847
Qy 660 AGGAAAGGGAAGTTTAAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAATTT 719
Db 848 AGGAAAGGGAAGTTTAAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAATTT 907
Qy 720 GGAATTTCAAGCGATTAATGATGAGTGAAGTGTGTTAGTCTGTTTGGGAGCAGGCTGC 779
Db 908 GGAATTTCAAGCGATTAATGATGAGTGAAGTGTGTTAGTCTGTTTGGGAGCAGGCTGC 967
Qy 780 TTTAGATCCATTCATATGTTTCAACTCTGCGCAACTCACTGACAGATTAATCATTTA 839
Db 968 TTTAGATCCATTCATATGTTTCAACTCTGCGCAACTCACTGACAGATTAATCATTTA 1027
Qy 840 CAATGGCTCATTTGACATCTCTCCCTGACACACACAGTTGATGCTGATTTTAAAGA 899
Db 1028 CAATGGCTCATTTGACATCTCTCCCTGACACACACAGTTGATGCTGATTTTAAAGA 1087
Qy 900 TACAGTTAGCATCTCTGAAAGCAGTTGGCTGTTTGTGAAAGTTCTTACAAATGCAACA 959
Db 1088 TACAGTTAGCATCTCTGAAAGCAGTTGGCTGTTTGTGAAAGTTCTTACAAATGCAACA 1147
Qy 960 ATCTGGTATGTCATGCTGATGAGTCTTACAAACAAATTTTCAGAGGACAGTACAA 1019
Db 1148 ATCTGGTATGTCATGCTGATGAGTCTTACAAACAAATTTTCAGAGGACAGTACAA 1207
Qy 1020 GTTCTCTAGACAGTGTCTCTCATACACTGGAAGGAGAGATTCATGAAGCAGTTTG 1079
Db 1208 GTTCTCTAGACAGTGTCTCTCATACACTGGAAGGAGAGATTCATGAAGCAGTTTG 1267
Qy 1080 TATTTTCAGAACCCAGAAATGTTTCAGGCTGACCCAGAGAAATTAATACAGCCTTTCTTTTAC 1139
Db 1268 CAGGTGAGAACCCAGAAATGTTTCAGGCTGACCCAGAGAAATTAATACAGCCTTTCTTTTAC 1327
Qy 1140 ATGGAAGAGCCTCGAGTGTGTTATGATACCATGATTCAGAGTTTGGATTTGATCA 1199
Db 1328 ATGGAAGAGCCTCGAGTGTGTTATGATACCATGATTCAGAGTTTGGATTTGATCA 1387
Qy 1200 GCGTTTGGATGAGAGGAGGACCAACCAAGCATGAAATTTTGGACAGATGCTATCAAGACT 1259
Db 1388 GCGTTTGGATGAGAGGAGGACCAACCAAGCATGAAATTTTGGACAGATGCTATCAAGACT 1447
Qy 1260 GGTAACTAT 1268
Db 1448 GGTGCTAT 1456

RESULT 9
CQ725791
LOCUS CQ725791 6807 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 11725 from Patent WO02068579.
ACCESSION CQ725791
VERSION CQ725791.1 GI:42287120
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Venter,C.J., Adams,M.C., Li,P.W. and Myers,B.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 11725 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1. .6807
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 36.1%; Score 1116.2; DB 6; Length 6807;
Best Local Similarity 99.7%; Pred. No. 6.4e-202;
Matches 1118; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 148 ATGCGAATCCTAAAGCGTTTCCTCGCTTGCATTCAGCTCCTCTGTGTTGGCGCTGGAT 207
Db 1 ATGCGAATCCTAAAGCGTTTCCTCGCTTGCATTCAGCTCCTCTGTGTTGGCGCTGGAT 60
Qy 208 TGGGCTAATGATATCTACAGACAACAGAGAAAATCTTGTGAGAGATTTGGCTGCTTAT 267
Db 61 TGGGCTAATGATATCTACAGACAACAGAGAAAATCTTGTGAGAGATTTGGCTGCTTAT 120
Qy 268 ACAGAGCAGTGAATCAAAAAATTTGGGAAAAGAAAATATCCAAATGTAATGAGCCAAAA 327
Db 121 ACAGAGCAGTGAATCAAAAAATTTGGGAAAAGAAAATATCCAAATGTAATGAGCCAAAA 180
Qy 328 CAATCTCTCATCAATATGATGAGATCTTACACAGTAAATGTAATCTTAAAGAACTT 387
Db 181 CAATCTCTCATCAATATGATGAGATCTTACACAGTAAATGTAATCTTAAAGAACTT 240
Qy 388 AAATTTTCAGGGTTGGGATAAAACATCATTTGGAACAACATTCATTCATAACACTGGGAAA 447
Db 241 AAATTTTCAGGGTTGGGATAAAACATCATTTGGAACAACATTCATTCATAACACTGGGAAA 300
Qy 448 ACAGTGAATTAATCTCACTAATGACTACCGTGTACGCGAGGAGTTTCAGAAATGCTG 507
Db 301 ACAGTGAATTAATCTCACTAATGACTACCGTGTACGCGAGGAGTTTCAGAAATGCTG 360
Qy 508 TTTAAAGCAGCAGATTAATCTTTTCACTGGGAAAATGCAATATGCTATCTGATGATCA 567
Db 361 TTTAAAGCAGCAGATTAATCTTTTCACTGGGAAAATGCAATATGCTATCTGATGATCA 420
Qy 568 GAGCATAGTTTGAAGACAAATAATTTTCCACTTCAGATGCAAAATCTACTGCTTTGATGCG 627
Db 421 GAGCATAGTTTGAAGACAAATAATTTTCCACTTCAGATGCAAAATCTACTGCTTTGATGCG 480
Qy 628 GACCGATTTTCAAGTTTGAAGAAAGCAGTCAAGGAAAAGGAAATTAAGAGCTTTATCC 687
Db 481 GACCGATTTTCAAGTTTGAAGAAAGCAGTCAAGGAAAAGGAAATTAAGAGCTTTATCC 540
Qy 688 ATTTTGTGAGGTTGGGACAGAGAAATTTTGCATTTCAAGCGATTTATGATGAGATC 747
Db 541 ATTTTGTGAGGTTGGGACAGAGAAATTTTGCATTTCAAGCGATTTATGATGAGATC 600
Qy 748 GAAAGTGTAGTCTGTTTGGGAGCAGCTGCTTTTAGATCCATTCATCTGTAACCTT 807
Db 601 GAAAGTGTAGTCTGTTTGGGAGCAGCTGCTTTTAGATCCATTCATCTGTAACCTT 660
Qy 808 CTGCCAAACTCAACTGACAAAGTATTAATTTAATGAGTCAATTTGACATCTCTCCCTGC 867
Db 661 CTGCCAAACTCAACTGACAAAGTATTAATTTAATGAGTCAATTTGACATCTCTCCCTGC 720
Qy 868 ACAGACAGTGTGATGCTGATTTTAAAGATACAGTATGATCTCTGGAAGCAGTTG 927
Db 721 ACAGACAGTGTGATGCTGATTTTAAAGATACAGTATGATCTCTGGAAGCAGTTG 780
Qy 928 GCTGTTTTTGTGGAAGTTCTTTTACAAATGCAAACTCTGCTTATGCTATGCTGATGACTAC 987

ORIGIN	/db_xref="taxon:10090"			
Query Match	30.3%;	Score 935.2;	DB 6;	Length 6887;
Best Local Similarity	86.4%;	Pred. No. 1.5e-167;		
Matches 1033;	Conservative 0;	Mismatches 163;	Indels 0;	Gaps 0;
QY	73	TTCTTTCGCTCCGCCCTCCCTCTCCATCTGTAGAAGCAGAGGAGCGCAGCGGAGGCC	132	
Db	31	TGCCTCGGTGTCCCAACCTCTGTACCTTAAGCTGTACCCCTGCGCTGGCGAGGGGCG	90	
QY	133	GCAGACCGTCTGGAATCGAATCTTAAGCGTTTCCTCGCTTGCATTCAGCTCTCTGT	192	
Db	91	CGACCGGCTGGAGATGCGAATCTCTGCGAGCTTCTCGCGTGTCTGCTCTCTGTGC	150	
QY	193	GTTCGCGCTGGAATGGGCTTAATGATATACAGACAACAGAGAAAATCTGTTGAAGAG	252	
Db	151	CTGTGCTGCTGAGCTGGCTTATGATATACAGACAACAGAGAAAATCTGTTGAAGAG	210	
QY	253	ATTGGCTGCTTATACAGAGCAGCTGAATCAAAAAAATTTGGGGAAGAAATATCCAA	312	
Db	211	ATTGGCTGCTTATACAGAGCAGCTGAATCAAAAAAATTTGGGGAAGAAATATCCAA	270	
QY	313	TGTAATAGCCCAAAACAATCTCTATCATATTTGATGAAGATCTTACACAAGTAATGTG	372	
Db	271	TGTAATAGCCCAAAAGCAGCTCTCTATTAATTTGATGAAGATCTTACACAAGTCAATGTG	330	
QY	373	AATCTTAAGAAACTTAATTTTCAGGGTTCGGATATAACATCATTTGGAACACACATTCATT	432	
Db	331	AATCTTAAGAAACTGAATTTTCAGGGTTCGGAAAGCGTCTTGGAAACACATTCATT	390	
QY	433	CATAACACTGGGAAAACAGTGGAAATTAATCTCACTAATGACTACCGTGTGAGCGGAGGA	492	
Db	391	CACAACACTGGGAAAACAGTGGAAATTAATCTCACTAATGACTACTCTCAGTGGAGGA	450	
QY	493	GTTCAGAAATGTGTTTAAAGCAGCAGATTAATCTTCTTCTGCTGGGAAAATGCAATATG	552	
Db	451	CTTTCAGAAAGGTCTTCAAGCGAAGCAGATTAATCTTCTTCTGCTGGGAAAATGCAATATG	510	
QY	553	TCATCTGATGGATCAGACATAGTTTGAAGGACAAAAATTTCCACTTGAGATCAAAATC	612	
Db	511	TCATCTGAGGATCGAAATAGCTTGAAGGACAGAGATTTCCACTGAGATCAAAATC	570	
QY	613	TACTGCTTTGATCGGACCGATTTTCAAGTTTGAAGGACAGCTCAAGGAAAAGGGAAG	672	
Db	571	TACTGCTTTGATCGGACAGATTTTCCAGTTTGAAGGACAGCTTAAAGGAAAAGGGAAG	630	
QY	673	TTAAGAGCTTATCCATTTTGTGAGGTGGGACAGAGAAAATTTGGATTTCAAGCG	732	
Db	631	TTAAGCGCTTTATCCATTTTATTTGAGGTGGAGTTGAAGAAAATTTGGATTAACAAGCC	690	
QY	733	ATTATTGATGGATCAAAAGTGTAGTCTTTTGGGAAGCAGGCTGCTTTAGATCCATTC	792	
Db	691	ATTATTGATGGAACTGAGAGTGTAGTCTTTTGGAAAGCAGGCTGCTTTAGATCCATTC	750	
QY	793	ATACTGTTGAACCTCTCGCAAACTCAACTGACAGATTAATCAATATGCTCATTTG	852	
Db	751	GTCTTCAGAACTCTCTCGCAAACTCAACTGACAGATTAATCAATATGCTCATTTG	810	
QY	853	ACATCTCTCTCGGACAGACAGATTTGATGCTGATTTTAAAGATACAGTTAGATC	912	
Db	811	ACATCTCTCTCGGACAGACAGATTTTAAAGATACAGTTAGATC	870	
QY	913	TCGAAAGCCAGTTGCTGTTTTGTGAAGTCTTCAATGCAACATCTGTTTATGTC	972	
Db	871	TCGAAAGCCAGTTGCTGTTTTGTGAAGTCTTCAATGCAACATCTGTTTATGTC	930	
QY	973	ATGCTGATGACTTACTTACAAAACAAATTTTCGAGAGCAACAGTCAAGTCTCTAGACAG	1032	
Db	931	ATGCTGATGACTTACTTACAAAACAAATTTTCGAGAGCAACAGTCAAGTCTCTAGACAG	990	
QY	1033	GTGTTTTCTCTATACACTCGAAAGGAGATTCATGAAGCAGTTTGTAGTTTCAGAACCA	1092	

AERDELSI HASLLVARGDLLEPSI VOYSDVASHQATTRAASDTLGRFSBAVFKYS
MVSQIESPRSDVWHVAVSGREPSYTVEGSHVTVTSYSSAMPLHGVSDVSDQSGLLI
NPISHPESSFTIPTASLLQPPALSGDEWSGASSDELLLPADGLRNLINISVP
SVAEFTYTVFADGIKPLSKSEMYGNETELKMSFSDMAYPSKSTVPRKMSDVHK
WSESLKETSVISKMSKVPSPSLVPTTKGPEQGVSHVPEIIPFQPTHTASQASGDT
WLKELGANSAAASDTSAREVHPSTOPLLYEAATPEANTEALLQPSQASDVTLK
TALPSVSDPILACTPQVEQSSSVSHRMASESSESMLHFTSPVILIDISPKVHST
PLQUTVPHSKSEKSEQLLASKSPQQVLPISLNFDEFQNAHDVDSQATPPKGRHAF
VTVPLSIDPEPTLNLKLVSESDIFPSSEISITDKVLGLPTLASDVLSTHDSVHAF
SGPISLTVMSPNRDSVTTAKLLPSTATSKLTQSARSADLGVGGSDGDDYDDDDYD
DIDGRFPVNMCMCLPVRESREKVMNDSDTQESSLVQSDPIIPLFPENTEBERNGT
GTVRVKSPSPMLPQNHNCKEDSDIQMSAVLPHTPGSKAMAVLSDRESGSGQT
SDSLNDNETSDFPDPVNEKDTGVLETDGTGAPGSPRSSTPSVTSIGSHGVNSSE
AG"

sig_peptide 106..177
mat_peptide 178..4941

/product="DSD-1-proteoglycan"

ORIGIN

Query Match 30.3%; Score 935.2; DB 10; Length 6887;
Best Local Similarity 86.4%; Pred. No. 1.5e-167;
Matches 1033; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 73 TTCCTTGCCTCCCTCCTCCTCACTCTGAGAAGCAGAGCGCGACGGCGAGGGGCC 132
DB 31 TGCCTCGTGTCCCACTCTGTACCTTAAGCTGTACCTCGGCTGGCGAGGGCGG 90

QY 133 GCAGACCGCTGGAAATCGGAATCCTAAAGCGTTTCCTCGCTTCGATTCAGCTCCTCTGT 192
DB 91 CGGACCGGGCTGGAGATCGGAATCCTGCAGAGCTTCCTCGCGTGGCTTCAGCTCCTGTGC 150

QY 193 GTTTCGCCCTCGAATGGCTTAATGGATCTACAGACAACAGAGAAATCTGTTGAAGAG 252
DB 151 CTGTGTCCCTCGACTGGCTGTATGGATCTACAGACAACAGAGAAATCTGTTGAAGAG 210

QY 253 ATTGCTCGTCTCTATACAGAGGAGCTGAATCAAAATAATTCGGAAGAAATATCCAA 312
DB 211 ATTGCTCGTCTCTACAGAGGAGCTGAATCAAAATAATTCGGAAGAAATATCCAA 270

QY 313 TGTAAATAGCCCAAAATCTCTATCAATATGATGAAGATCTTACCAAGAAATATG 372
DB 271 TGTAAATAGCCCAAGCAGTCTCTATTAATTAATGATGAAGATCTTACCAAGTCAATGTG 330

QY 373 AATCTTAGAAACTTAATTTTCAGGGTTGGATTAATCAATCATTTGGAAACACATTCAT 432
DB 331 AATCTTAGAAACTTAATTTTCAGGGTTGGAAAGCGTCTTGGAAACACGTTCAAT 390

QY 433 CATACACTGGGAAACAGTGGAAATTAATCTCACTAATGACTACCGTGTGAGCGGAGGA 492
DB 391 CACACACTGGGAAACAGTGGAAATTAATCTCACTAATGACTATCTCAGTGGAGGA 450

QY 493 GTTTCAGAAATGGTGTGTTAAAGCAAGCAAGATACTTTCTCGGGGAAATGCAATATG 552
DB 451 CTTTCAGAAAGGTCTTCAAGCGCAAGCAAGATACTTTCCACTGGGAAATGCAATGTG 510

QY 553 TCATCTGATGATCAGACATAGTTTAGAAGGACAAATAATTTCCACTTGAGATGCAATC 612
DB 511 TCATCTGAGGATCGGAACATAGCTTAGAAGGACAAAGATTTCCCACTGGATGCAAGTC 570

QY 613 TACTGCTTTGATGCGGACCGAATTTTCAAGTTTGTAGGAAGCAGTCAAAAGGAAAGGGAAG 672
DB 571 TACTGCTTTGATGCGGACAGATTTTCCAGTTTGTAGGAAGCAGTTAAAGGAAAGGAGA 630

QY 673 TTAAGAGCTTTATCAATTTTGTGAGGTGGGACAGAAATAATTTGATTTCAAAGCG 732
DB 631 TTAAGCGCTTTATCAATTTTATTTGAGGTGGGTTGAAGAAATTTGATTTCAAAGGCC 690

QY 733 ATTATTTGATGAGTTCGAAAGTGTAGTGTGTTTGGGAAGCAGGCTGCTTTAGATCCATTC 792
DB 691 ATTATTTGATGAGTTCGAGAGTGTAGTGTGTTTGGGAAGCAGGCTGCTTTAGATCCATTC 750

QY 793 ATACTGTTGAACCTTCTGCAAACTCAACTGACAGATATTACATTTTACAATGGCTCAATG 852

DB 751 GTCTTGAGAAACCTCCTCGCAAACTCCACGTGACAAGTATTACATTTTACAATGGATCATTTG 810

QY 853 ACATCTCCTCCTCGCAAGACACAGATTTGACTGAGTGTGTTTAAAGATACAGATTAGCATC 912

DB 811 ACATCCCTCCTCGCAAGACACACCGTGGAAATGGATTCTTTTAAAGGATACAGATTAGCATC 870

QY 913 TCTGAAAGCCAGTTGGCTGTTTGTGGAAGTCTTCAATGCAACAACTCTGTTTATGTC 972

DB 871 TCTGAAAGCCAGTGGCTGTATTGGAAGTCTTCAATGCAACAGTCTGGGTATGTC 930

QY 973 ATGCTGATGACTACTTACAAAACAAATTTTCGAGAGCAACAGTACAAAGTCTCTAGACAG 1032

DB 931 ATGTTGATGATTTTACAAAACAAATTTTCGAGAGCAACAGTACAAAGTCTTCCAGGAG 990

QY 1033 GTGTTTTCTCATACACTTGGAAAGAAAGATTCATGCAAGCAGTTTGTAGTTTCAAGACCA 1092

DB 991 GTGTTTTCTCATATATCGAAAGAAAGATTCACCAAGTAGTGTGTAGTTTCAAGACCA 1050

QY 1093 GAAATGCTTCAGGCTGACCCAGAGAAATTTATACAGAGCTTCTGTTTATACATGGAAGACCT 1152

DB 1051 GAAATGTCAGAGCTGACCTTGAGAAATTTACCAAGCTTCTGGTCACTGGAAGACCT 1110

QY 1153 CGAGTCTGTTTATGATACCATGATTGAGAAGTTTGACGTTTTGTACCAAGCAGTTGGATGGA 1212

DB 1111 CGGTCGTTTATGAGCCCATGATTGAGAAGTTTGACGTTCTGTACCAAGCAGCTCGCGGA 1170

QY 1213 GAGGACCAACCAAGCATGAAATTTTTCAGAGATGGCTTATCAAGACTTGGTAACTAT 1268

DB 1171 AATGACCAAGCAGCATGAGTTCTTAACAGATGGCTTATCAGGACTTGGGTGCCAT 1226

RESULT 13

AX683146
LOCUS
DEFINITION Sequence 120 from Patent EP1279744.
ACCESSION AX683146
VERSION AX683146.1 GI:29370167
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Rattus.
REFERENCE 1
AUTHORS Brookebank, R.A., Dixon, A.K., Lee, K. and Pincock, R.D.
TITLE Identification and use of molecules implicated in pain
JOURNAL Patent: EP 1279744-A 120 29-JAN-2003;
WARNER-LAMBERT COMPANY (US)
FEATURES
Location/Qualifiers
source 1..6801
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/note="Phosphacan"

ORIGIN

Query Match 30.0%; Score 926.4; DB 6; Length 6801;
Best Local Similarity 86.4%; Pred. No. 7.2e-166;
Matches 1023; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 85 CCCTCCCTCTCCACTCTGAGAAGCAGAGGAGCGCGACGGCGAGGGCGCGAGACCGTCTG 144

DB 43 CCCACTCTTCGACCCCTTAAGCTGTACCCCTGGCGTGGCGAGGGCGCGGACCGCGCTG 102

QY 145 GAAATGCGAATCTTAAAGCGTTTCTCGTTGCAATTCAGTCTCTCTGTGTTTGGCGCTG 204

DB 103 GAGATGCGAATCTTCGACAGCTTCTCGCGTGGCTTCAAGTCTCTGTGCTGTGCTGCTG 162

QY 205 GATTGGGCTTAATGATACTACAGACACAGAAACTTGTGGAAGATTTGGCTGTGCTG 264

DB 163 GACTGGGCTTATGATATCTACAGACAAAGAAACTTGTGGAAGATTTGGCTGTGCTG 222

QY 265 TATACAGGAGCAGTGAATCAAAAAAATTGGGAAAGAAATATATCAACATGTAATAGCCCA 324

RESULT 15	
RNU09357	
LOCUS	7851 bp mRNA linear ROD 11-JUL-1995
DEFINITION	Rattus norvegicus receptor-type protein tyrosine phosphatase zeta/beta mRNA, complete cds.
ACCESSION	U09357
VERSION	U09357.1 GI:487780
KEYWORDS	
SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus
REFERENCE	Maurel P., Meyer-Puttlitz B., Flad M., Margolis R.U. and Margolis R.K. Nucleotide sequence and molecular variants of rat receptor-type protein tyrosine phosphatase-zeta/beta DNA Seq. 5 (5), 323-328 (1995) 96063026 PUBMED 7579589
AUTHORS	1 (bases 1 to 7851) 2 (bases 1 to 4951) Maurel P., Rauch U., Flad M., Margolis R.K. and Margolis R.U. Phosphacan, a chondroitin sulfate proteoglycan of brain that interacts with neurons and neural cell-adhesion molecules, is an extracellular variant of a receptor-type protein tyrosine phosphatase Proc. Natl. Acad. Sci. U.S.A. 91 (7), 2512-2516 (1994) 94159772 PUBMED 7511813
JOURNAL	3 (bases 1 to 7851) Margolis R.U. Direct Submission Submitted (02-MAY-1994) Richard U. Margolis, Pharmacology, New York University Medical Center, 550 First Avenue, New York, NY 10016, USA
MEDLINE	
PUBMED	
FEATURES	Location/Qualifiers 1..7851 /organism="Rattus norvegicus" /mol_type="mRNA" /strain="Sprague Dawley" /db_xref="taxon:10116" /clone="c12, c21, pBSP3, pBSP4, pBcyt6alpha" /sex="male" /tissue type="brain" /clone_lib="oligo-dt primed rat brain cdna library in Lambda ZAPII from Stratagene" /dev_stages="6 weeks" <1..105 106..7056 /note="full length rPTP zeta/beta, including phosphacan (3F8 proteoglycan); for the extracellular variant (phosphacan, 3F8 proteoglycan) see GenBank Accession Number U04998; rat homolog of the human receptor-type protein tyrosine phosphatase zeta/beta, Swiss-prot accession number P23471, PIR accession number A46700" /citation=[2] /codon_start=1 /product="receptor-type protein tyrosine phosphatase zeta/beta" /protein_id="AAC52207.1" /translation="MRLIQSLFCVQLLCVLDWVGYYRQKLVRBIGWSYTGAL NOKWKGKYPICNSPKQSINIDEDLTQVNVLKKLPQGWKEPSLENTPIHNTGTV EINIUNTYLISGLSEKVPFKASMTFWGKNLVSSSGEHSLEGKFPLEMIQICYFDA DRFSSETVNGKGRRLALSILFEIGVECNLDYKAIDGTESVRSFGQAALDPFLCA NLIPNSTDKYIIYNGSTLTPSPDTVEWIVPKDTSVSISQAVFCENVLTMOOSGVVM LMDYLQNRFQQKQVFSSVTGKEIHAEVCSSPEPNQADPENYTSLLITWER PRVVLYTMIEKFALVLPKGNQDTHPELTDCGYQDLGAILNNLIPNMSYVLQIVAI C SNIYLGKYSDLI VDMTEBAEDLDFELIGTBEI I KEENYGKHNEBDTLGNPGRDSA TNOIRKCEPVSTITHNHMGTKYNEAKTNRPTRSFSGSKSDVLNLSNFTSQOVA EFNREREMSPSQIGTLPPHPVSSAINSGSKTLLVFPQMNISGTAESLNMYSTI PEYKVSADLGEENLTDPLKDSGADDSSGSSPASSTVPFSDNLSHGYTSSSDPTFA
5'UTR	
CDS	

QY 325 AAACAATCTCCTATCAATATTTGATGAAGATCTTACACAAGTAAATGTGAATCTTTAAGAAA 384
|||
Db |||
QY 283 AAGCAGTCTCTATTTAATATTTGATGAAGATCTTACACAAGTAAATGTGAATCTTTAAGAAA 342
|||
Db |||
QY 385 CTTAAATTTCCAGGTTGGGATAAACATCATTTGGAAAAACACATTCATTATATAACACTGGG 444
|||
Db |||
QY 343 CTGAAATTTCCAGGTTGGGAAAAACCATCTCTGGAAAAACAGTTTATTTCCACACACTGGG 402
|||
Db |||
QY 445 AAAACAGTGGAAATTAATCTCACTAATGACTACCGTGTACGGGAGGAGTTTCAGAAATG 504
|||
Db |||
QY 403 AAAACAGTGGAAATTAATCTCACTAATGACTACTATCTCAGTGGAGGACTTTTCAGAAAAG 462
|||
Db |||
QY 505 GTGTTTAAAGCAAGCAAGATAACTTTTCCACTGGGAAAAATGCAATATGTCTATCTGATGGA 564
|||
Db |||
QY 463 GTCTTCAAGGCAAGCAAAATGACTTTTCTCTGGGAAAAATGCAATGTCTATCGGAAGGA 522
|||
Db |||
QY 565 TCAGAGCATAGTTTAGAAGGACAAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGAT 624
|||
Db |||
QY 523 TCGGAACACAGCTTAGAAGGCGAGAAGTTTCCACTCGAGATGCAAAATCTACTGCTTCGAT 582
|||
Db |||
QY 625 CGGACCGAATTTCAAGTTTGGAGAGCAGTCAAGGAAAAAGGAAAGTTAAGAGCTTTTA 684
|||
Db |||
QY 583 CGGACCGAATTTTCAAGTTTGGAGAAAACAGTTAAGGAAAAAGGAGTTAAGGGCTCTA 642
|||
Db |||
QY 685 TCCATTTTGTGTTGAGTTGGGACAGAAAAATTTGGATTTCAAAGCGATTTATTGATGGA 744
|||
Db |||
QY 643 TCCATTTTGTGATTTGGAGTTGGAAGAAAATTTGGATTTCAAAGCGATTTATTGATGGG 702
|||
Db |||
QY 745 GTCGAAAGTGTAGTTCGTTTGGGAAGCAGGCTGCTTTAGATCCATTCTATCTGTTGAAC 804
|||
Db |||
QY 703 ACTGAAAGTGTAGTTCGTTTGGGAAGCAGGCTGCTTTAGATCCGTTCTATCTTCGAGAAC 762
|||
Db |||
QY 805 CTTCTGCCAAACTCACTGACAGATTTATTAATTTACAATGGCTCATTGACATCTCCTCCC 864
|||
Db |||
QY 763 CTCCTGCCAAACTCACTGACAGATTTATTAATTTACAACGGATCACTGACATCCCTCCC 822
|||
Db |||
QY 865 TGCACAGACACAGTTGACTGATTTGTTTAAAGATACAGTTAGCATCTCTGAAAGCCAG 924
|||
Db |||
QY 823 TGCACAGACACCGTCGAATGATTTGTTTAAAGATACAGTTAGCATCTCTGAAAGCCAG 882
|||
Db |||
QY 925 TTGCTGTTTTTGTGAAGTTCTTACAATGCAACAATCTGGTTATGTCAATCTGATGGAC 984
|||
Db |||
QY 883 CTGGCTGTATTTTGTGAAGTTCTCACAATGCAACAGTCCGGGTATGTCAATGTGATGGAT 942
|||
Db |||
QY 985 TACTTACAAAACAATTTTCGAGAGCAACAGTACAGTTCTCTAGACAGGTTGTTTCCTCA 1044
|||
Db |||
QY 943 TACTTACAAAACAATTTTCGAGAAACAACAGTACAGTTTTCAGGCGAGGTGTTTCTTCG 1002
|||
Db |||
QY 1045 TACACTGGAAGGAAAGAGATTTCAGAACAGTTTGTAGTTCCAGAACCAAGAAAATGTTCCAG 1104
|||
Db |||
QY 1003 TACACTGGGAAGGAAGAGATTCTAGNAGCAGTTGCGATTTCAGAACCAAGAAAATGTGCA 1062
|||
Db |||
QY 1105 GCTGACCCAGAGAAATTAACAGCCTTCTGTATCATGGGAAAAGACCTTCGAGTCGTTTAT 1164
|||
Db |||
QY 1063 GCTGACCCCTGAGAAATTAACAGCCTTCTGTATCATGGGAAAAGCCTTCGGGTTGTTTAT 1122
|||
Db |||
QY 1165 GATACCATGATGAGAAATTTGAGTTTGTACCGAGAGTTGGATGGAGAGACCAAAACC 1224
|||
Db |||
QY 1123 GACACAATGATGAGAAATTTGCGGTCTGTACCGCCACTGGAGGGGAAACGACCAAAACC 1182
|||
Db |||
QY 1225 AAGCATGAATTTTTCACAGATGGCTATCAAGACTTGGTAACTAT 1268
|||
Db |||
QY 1183 AAGCATGAGTTTTTACAGATGGCTATCAGGACTTGGGTGCCAT 1226
|||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 11:30:24 ; Search time 9626 seconds
(without alignments)
12222.798 Million cell updates/sec

Title: US-09-983-000a-1
Perfect score: 3091
Sequence: 1 cacacatgcgcgcacgat.....ttatgtagtaagtgtatata 3091

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_ges1:*
9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	940	30.4	1614	3	AK081698	AK081698 Mus musculus
2	677.8	21.9	755	4	BG911384	BG911384 602807687
3	623.8	20.2	757	6	CD350667	CD350667 UI-M-GV0-
4	562.4	18.2	572	7	CN418533	CN418533 170004243
5	512.4	16.6	572	2	BF365490	BF365490 KC1-NT003
6	498	16.1	880	4	BG202130	BG202130 RST21484
7	494.8	16.0	701	7	CN383698	CN383698 4117208 B
8	484.2	15.7	585	6	CA879683	CA879683 K0976A09-
9	481.6	15.6	718	7	CK781524	CK781524 UI-M-GV0-
10	475.4	15.4	695	1	AL706826	AL706826 DKF2p686J
11	473.2	15.3	664	6	CA874874	CA874874 K0935E06-
12	453.6	14.7	780	6	CD803608	CD803608 UI-M-GV0-
13	430.8	13.9	522	6	CA870557	CA870557 K0903A03-
14	414.6	13.4	629	5	BX642570	BX642570 DKF2p686L
15	409	13.2	498	6	CA872941	CA872941 K0921G05-
16	401	13.0	522	6	CA871074	CA871074 K0907F04-
17	388.8	12.6	564	7	CK344902	CK344902 K0977H04-
18	386.8	12.5	598	1	AL706550	AL706550 DKF2p686J
19	382.8	12.4	461	6	CD803995	CD803995 UI-M-GV0-
20	376.2	12.2	552	7	CV040479	CV040479 4138002 B
21	373.2	12.1	450	6	CA878472	CA878472 K0967A01-
22	371	12.0	681	7	CF744132	CF744132 UI-M-GV0-
23	360.4	11.7	583	5	BP230191	BP230191
24	355.6	11.5	741	7	CF744324	CF744324 UI-M-GV0-

25	346	11.2	514	1	AL712835	AL712835 DKF2p686H
26	338.6	11.0	499	6	CA877556	CA877556 K0957F09-
27	336.6	10.9	565	7	CF746133	CF746133 UI-M-GV0-
28	326.2	10.6	940	5	BU913714	BU913714 AGENCOURT
29	324.4	10.5	517	1	AL584698	AL584698 AL584698
30	324	10.5	518	2	BH531767	BH531767 601230891
31	319.6	10.3	582	5	BF230899	BF230899 BP230899
32	312.6	10.1	522	7	CF538636	CF538636 UI-M-GV0-
33	298.4	9.7	733	5	BU317596	BU317596 603405758
34	296.4	9.6	298	7	N64320	N64320 YZ46C11-81
35	288.4	9.3	343	2	AW274814	AW274814 XM61a06.X
36	288.4	9.3	355	2	AW205697	AW205697 UI-H-B11-
37	281.4	9.1	336	1	AI203217	AI203217 Qr29a04.X
38	277.6	9.0	338	1	AI340237	AI340237 Qx86a06.X
39	275.8	8.9	327	2	BF471307	BF471307 UI-M-BH3-
40	267.2	8.6	580	5	BP229566	BP229566 BP229566
41	265.2	8.6	300	1	AI867704	AI867704 WB37C02.X
42	262.4	8.5	281	1	AI961037	AI961037 WQ58E03.X
43	261.8	8.5	587	6	CA888172	CA888172 B0143C06-
44	254.2	8.2	784	7	CO423210	CO423210 GGE2HT101
45	252	8.2	539	1	AL910230	AL910230 AL910230

ALIGNMENTS

RESULT 1	AK081698	1614 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	Mus musculus	16 days embryo head cDNA, RIKEN full-length enriched library, clone: C130068O22	product: protein tyrosine phosphatase, receptor type, Z, full insert sequence.		
DEFINITION	AK081698	AK081698.1	GI:26349318		
ACCESSION	AK081698	HTC; CAP trapper.			
VERSION	AK081698	Mus musculus (house mouse)			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus				
ORGANISM	Mus musculus				
REFERENCE	1	Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watanabe, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
AUTHORS	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multispillar sequencer				
TITLE	Genome Res. 10 (11), 1757-1771 (2000)				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4	The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.			
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409, 685-690 (2001)				
PUBMED	11076861				
REFERENCE	5	The PANTOM Consortium and the RIKEN Genome Exploration Research			

TITLE JOURNAL REFERENCE AUTHORS	Group Phase I & II Team.									
	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs									
	Nature 420, 563-573 (2002)									
	6 (bases 1 to 1614)									
FEATURES	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.									
	Direct Submission									
	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)									
	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.									
	Please visit our web site for further details.									
	URL:http://genome.gsc.riken.jp/									
	URL:http://fantom.gsc.riken.jp/									
	Location/Qualifiers									
	1. .1614									
	/organism="Mus musculus"									
/mol_type="mRNA"										
/strain="C57BL/6J"										
/db_xref="FANTOM_DB:C130068022"										
/db_xref="taxon:10090"										
/clone="C130068022"										
/tissue_type="head"										
/clone_lib="RIKEN full-length enriched mouse cDNA library"										
/dev_stage="16 days embryo"										
327. _>1613										
CDS	/note="unnamed protein product; protein tyrosine phosphatase, receptor type, z (MGD MGI:97816, GB NM_011219, evidence: BLASTN, 99%, match=1393) putative"									
	/codon_start=1									
ORIGIN	/protein_id="BAC38299.1"									
	/db_xref="GI:26349319"									
Query Match	/translation="MRILQSFACVQLLCLRLDWAYGYTRQRLVVEIGWSYTGAL NQKWKGYPTCNPKQSPINIDLTQVNVNKLKLPQGWKASLENTFIHNTKTV EINTNDYLSGSLSEKVPKSKITFHMKCNVSSSEGSLSLEQKPLEMOVYCPDA DRFSPEAVAGKRLRLSLPFGVEENLDYKALIDGTESVRFKQALDPPVLQ NLLPSTDKYIINGSLTSPCTDVEWVFQTVSISEQLAVFCEVILMQQGGYVM LMDYQNNFRQQYKFSRQYSSYTGKEEIHVEVCSSEPNVQADPNYSLLVTSR PRVYVDAMIEKFAVLYQPLAGNDQAKHFLFDGYQDLGAILNLLPNNYSVYLQVAVC SNGLYGKYSQDLIVDMPTDEALDLPPELIGTEBIK"									
	30.4%; Score 940; DB 3; Length 1614;									
Best Local Similarity	86.6%; Pred. No. 3.9e-189;									
	Matches 1036; Conservative 0; Mismatches 160; Indels 0; Gaps 0;									
Qy	73	TTCTTCCTCCCTCCCTCTCCACTCTCAGAGCAGAGCGCGCGCGGGGCC	132							
Db	252	TGCTCGGTGTCCCACTCTGTACCTTAAGCTGTACCTCGGCTGGCGAGGGCCG	311							
Qy	133	GCAGACCGTCTGGAATCGGAATCTTAAAGCGTTTCTCGCTTGCAATTCAGCTCCTCTGT	192							
Db	312	CGGACCGGCTGGAGATCGGAATCTCGCAGAGCTTCTCGCGTTCGTTTCACTCCTGTGC	371							


```

RESULT 2
BG911384
LOCUS      BG911384      755 bp      mRNA      linear      EST 05-JUN-2001
DEFINITION 602807687F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4939965
5', mRNA sequence.
ACCESSION  BG911384
VERSION     BG911384.1  GI:14291860
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 755)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/ILLNL at:
            http://image.llnl.gov
            Plate: LLAM10877 row: m column: 22
            High quality sequence stop: 747.

FEATURES             Location/Qualifiers
     source           1..755
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4939965"
                     /tissue_type="anaplastic oligodendroglioma with lp/19q
                     loss"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI CGAP_Brn67"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
                     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 2.3 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match      21.9%; Score 677.8; DB 4; Length 755;
Best Local Similarity 98.3%; Pred. No. 2.2e-133;
Matches 748; Conservative 0; Mismatches 7; Indels 6; Gaps 6;

QY 300 GAAATATCCACATGTAATAGCCCAAAACAATCTCTATCAATATTGATGAAGATCTTTAC 359
DB 1 GAAATATCCACATGTAATAGCCCAAAACAATCTCTATCAATA-TGATGAAGATCTTAC 59
QY 360 ACNAGTAAATGTAATCTTAAGAAACTTAAATTTGAGGTTGGGATAAAACATCATTTGGA 419
DB 60 ACNAGTAAATGTAATCTTAAGAAACTTAAATTTGAGGTTGGGATAAAACATCA-TGGA 118
QY 420 AAACACATTTCATTATAACACTGGGAAACAGTGAATTAATCTCACTAATGATCAACCG 479
DB 119 AAACACATTTCATTATAACACTGGGAAACAGTGAATTAATCTCACTAATGATCAACCG 178
QY 480 TGTTCAGCGGAGGAGTTTCAGAAATGGTGTTTAAAGCAAGCAAGATAAATCTTTCACTGGGG 539
DB 179 TGTTCAGCGGAGGAGTTTCAGAAATGGTGTTTAAAGCAAGCAAGATAAATCTTTCACTGGGG 238
QY 540 AAATATGCAATATGTCATCTGATGGATCAGAGCATAGTTTAGAGGACAAAATTTCCACT 599
DB 239 AAATATGCAATATGTCATCTGATGGATCAGAGCATAGTTTAGAGGACAAAATTTCCACT 298
QY 600 TGAGATGCAATCTACTCTTTGATGCGGACCGATTTCAGATTTTAGAGGAGCAGTCAA 659
DB 299 TGAGATGCAATCTACTCTTTGATGCGGACCGATTTCAGATTTTAGAGGAGCAGTCAA 357
QY 660 AGGAAAGGGAAGTTAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAATTT 719

```

```

Db 358 AGGAAAGGGAAGTTAAGAGCTTTATCCA-TTTGTTGAGGTTGGGACAGAGAAATTT 416
QY 720 GGATTTCAAGCGAATTATTGATGAGTGAAGTGTAGTGTGTTTGGGAGCAGGCTGC 779
Db 417 GGATTTCAAGCGAATTATTGATGAGTGAAGTGTAGTGTGTTTGGGAGCAGGCTGC 476
QY 780 TTTAGATCCATTTCATCTGTTGAACCTTCTGCCAACTCAACTGCAAGTATTACATTTA 839
Db 477 TTTAGATCCATTTCATCTGTTGAACCTTCTGCCAACTCAACTGCAAGTATTACATTTA 536
QY 840 CAATGGCTCATTTGACATCTCTCCCTGCACAGACACAGTTGATGTGTTTAAAGA 899
Db 537 CAATGGCTCATTTGACATCTCTCCCTGCACAGACACAGTTGATGTGTTTAAAGA 596
QY 900 TACAGTTAGCATCTCTGAAGCCAGTTGGCTGTTTTTGTGAAGTTCTTACATGCAACA 959
Db 597 TACAGTTAGCATCTCTGAAGCCAGTTGGCTG-TTTATGTGAAGTTCTTACATGCAACA 655
QY 960 ATCTGGTTATCTCATCTGATGGACTACTTACAAAACAATTTTCGAGAGCAACAGTACAA 1019
Db 656 ATCTGGTTATCTCATCTGATGGACTACTTACAAAAGAATTTTCGAGAGCAACAGTACAA 715
QY 1020 GTTCTCTAGACAGGTGTTTCTCTCATACACTCGAAAGGAAG 1060
Db 716 GTTCTCTAGACAGGTGTTT-CTCATACACTGGACAGGAAG 755

RESULT 3
CD350667
LOCUS      CD350667      757 bp      mRNA      linear      EST 09-JUL-2003
DEFINITION UI-M-GIO-c9g-f-21-0-UI-r1 NIH_BMAP_G10 Mus musculus cDNA clone
IMAGE:6852982 5', mRNA sequence.
ACCESSION  CD350667
VERSION     CD350667.1  GI:31142182
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 757)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mouseefi.html
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)

Seq primer: pyX-5.
            Location/Qualifiers
     source           1..757
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="NIH_BMAP_G10"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_G10"
                     /note="Organ: Brain; Vector: pyX- Asc; Site 1: Ecor I;
                     Site 2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated

```

with EcoR I adaptor, digested with NotI and then cloned directionally into pX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BNAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

```
Query Match      20.2%; Score 623.6; DB 6; Length 757;
Best Local Similarity 89.6%; Pred. No. 6.7e-122;
Matches 670; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 317 ATAGCCCAAAACATCTCTATCAATATGTAGTAAGATCTTTACACAGTAATAATGTGAATC 376
DB 1 ATAGCCCAAAACAGTCTCTATTAATATGTAGTAAGATCTTTACACAGTCAATGTGAATC 60
QY 377 TTAGAAAATCTTAAATTTTCAGGTTGGGATAAAACATCATTCGMAAAACACATTCATTCATA 436
DB 61 TTAGAAAATCTGAATTTTCAGGTTGGGAAAAGCGTCTTGGAAAACACGTTTCATTCACA 120
QY 437 ACCTGGGAAAACAGTGGAAATTAATCTCACTAATGACTACCGTGTGACGGGAGGAGTTT 496
DB 121 ACCTGGGAAAACAGTGGAAATTAATCTCACTAATGACTACTATCTCAAGTGGAGACTTT 180
QY 497 CAGAAATGGTGTAAACCAAGCAAGATAACTTTTCACTCTGGGAAAATGCAATATGTCAT 556
DB 181 CAGAAATGGTCTTCAAGCAAGCAAGATAACTTTTCACTCTGGGAAAATGCAATATGTCAT 240
QY 557 CTGATGATCAGACATAGTTTGAAGGACAAAAATTTCCACTTGGATGCAAAATCTACT 616
DB 241 CTGAGGATCGAACATAGCTTAGAGGACAGAGTTCCCACTGGAGATGCAATCTACT 300
QY 617 GCTTTGATCGGACCGATTTTCAAGTTTGAAGGACAGTCAAGGAAAAGGAAAGTTAA 676
DB 301 GCTTTGATCGGACAGATTTTCCAGTTTGAAGGACAGTCTTAAAGGAAAAGGAAAGTTAA 360
QY 677 GAGCTTTATCAATTTGTTGAGGTTGGGACAGAGAAATTTGATTTCAAAGCGATTA 736
DB 361 GGGCTTTATCCATTTTATTTGAGGTTGGAGTTGAAGAAAATTTGGATTTACAAAGCCATTA 420
QY 737 TTGATGAGTGGAAAGTGTAGTCTGTTTGGAGAGCAGGCTGCTTTAGATCCATTTCTATAC 796
DB 421 TTGATGGAACCTGAGAGTGTAGTCTGTTTGGAGAGCGCTGCTTTAGATCCATTTCTGCT 480
QY 797 TGTGAACCTTTCTGCCAACTCAACTGACAAAGTATTAATTAACAATGGCTCATTTGACAT 856
DB 481 TGCAGAACCTCTCTGCCAACTCCACTGACAAAGTATTAATTAACAATGGATCATTTGACAT 540
QY 857 CTCCTCCCTGCACAGACAGTGTGATGCTGATGTTTAAAGATACAGTTAGCATCTCTG 916
DB 541 CCCCTCCCTGCACAGACACCGTGGAAATGGATTTTAAAGGATACAGTTAGCATCTCTG 600
QY 917 AAAAGCCAGTGGCTGTTTGTGAGGTTCTTACAAATGCAACAATCTGTTTATGTCATGC 976
DB 601 AAAGCCAGTGGCTGTTTGTGAGGTTCTTACAAATGCAACAAGTCTGGGTATGTCATGT 660
QY 977 TGATGGACTACTTACAAACAATTTTCGAGAGCAACAGTACAAGTTCTCTAGACAGGTTG 1036
DB 661 TGATGGATTACTTACANAACAATTTCCGAGAACACAGTACAAGTTTTCAGGCGAGGTTG 720
QY 1037 TTTCTCATACACTGGAAAGGAGAGAT 1064
DB 721 TTTCTCATATACTGGAAAGGAGAGAT 748
```

RESULT 4

```
CN418533 LOCUS CN418533 572 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000424334377 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN418533
VERSION CN418533.1 GI:47406127
```

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

1 (bases 1 to 572)

Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,

Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,

Lebkowski, J. and Stanton, L.W.

Transcriptome characterization elucidates signaling networks that

control human ES cell growth and differentiation

Nat. Biotechnol. 22 (6), 707-716 (2004)

Contact: Brandenberger R

Regenerative Medicine

Geron Corporation

230 Constitution Drive, Menlo Park, CA 94025, USA

Tel: 650 473 8658

Fax: 650 473 7760

Email: rbrandenberger@geron.com

Insert Length: 572 Std Error: 0.00.

Location/Qualifiers

1. 572

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="embryonic stem cells, cell lines H1, H7, and

H9"

/clone_lib="GRN_ES"

/notes="oligo dT primed, full-length enriched cDNA library

from undifferentiated hES cell lines H1 (p32), H7 (p29),

and H9 (p26) maintained in feeder-free conditions"

ORIGIN

```
Query Match      18.2%; Score 562.4; DB 7; Length 572;
Best Local Similarity 99.8%; Pred. No. 7.5e-109;
Matches 563; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 234 GAGAAACTTTTGAAGAGATTCGCTGCTCTATACAGGAGCAGTGAATCAAAAAATG 293
DB 9 GAGAAACTTTTGAAGAGATTCGCTGCTCTATACAGGAGCAGTGAATCAAAAAATG 68
QY 294 GGGAAAGAAATATCCAAATGTAATAGCCCAAAACAAATCTCTATCAATATTCATGA 353
DB 69 GGGAAAGAAATATCCAAATGTAATAGCCCAAAACAAATCTCTATCAATATTCATGA 128
QY 354 TCTTACACAAATGTAATGTAATCTTAAGAACTTAAGAACTTAAGAACTTAAGAACT 413
DB 129 TCTTACACAAATGTAATGTAATCTTAAGAACTTAAGAACTTAAGAACTTAAGAACT 188
QY 414 ATTGGAAAACACATTCATTCATTAACACTGGGAAAACAGTGGAAATTAATCTCACTA 473
DB 189 ATTGGAAAACACATTCATTCATTAACACTGGGAAAACAGTGGAAATTAATCTCACTA 248
QY 474 CTACCGTGTGACGGGAGGAGTTTCAGAAAATGGTGTTTAAAGCAAGCAAGATACTTTTCA 533
DB 249 CTACCGTGTGACGGGAGGAGTTTCAGAAAATGGTGTTTAAAGCAAGCAAGATACTTTTCA 308
QY 534 CTGGGAAAAATGCAATATGTCATCTGATGATCAGACATAGTTTGAAGACAAAAATT 593
DB 309 CTGGGAAAAATGCAATATGTCATCTGATGATCAGACATAGTTTGAAGACAAAAATT 368
QY 594 TCCACTTTGAGATGCAAAATCTACTGCTTTGATGGGACCGGATTTTCAAGTTTGAAGGAGC 653
DB 369 TCCACTTTGAGATGCAAAATCTACTGCTTTGATGGGACCGGATTTTCAAGTTTGAAGGAGC 428
QY 654 AGTCAAAAGGAAAAGGGAAGTTAAGAGCTTTATCCATTTTGTGTTGAGGTTGGGACAGAAGA 713
DB 429 AGTCAAAAGGAAAAGGGAAGTTAAGAGCTTTATCCATTTTGTGTTGAGGTTGGGACAGAAGA 488
QY 714 AAATTTGGATTTCAAACGATTTATGATGAGTGCAGAAAGTGTAGTCTGTTTGGGAAGCA 773
DB 489 AAATTTGGATTTCAAACGATTTATGATGAGTGCAGAAAGTGTAGTCTGTTTGGGAAGCA 548
```

```
QY 774 GCTGCTTTAGATCCATTCATACT 797
Db 549 GGTGCTTTAGATCCATTCATACT 572

RESULT 5
BF365490/c
LOCUS BF365490 572 bp mRNA linear EST 24-NOV-2000
DEFINITION RC1-NT0033-090800-016-h04 NT0033 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF365490
VERSION BF365490.1 GI:11327515
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 572)
Dias Neto,B., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,P., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=RC1&t2=RC1-NT0033-
090800-016-h04&t3=2000-08-09&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 48
High quality sequence stop: 572.
FEATURES
source
1..572
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NT0033"
/note="Organ: nervous tumor; Vector: puc18; Site 1: Smal;
Site 2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
ORIGIN
Query Match 16.6%; Score 512.4; DB 2; Length 572;
Best Local Similarity 96.8%; Pred. No. 3.3e-98;
Matches 544; Conservative 0; Mismatches 16; Indels 2; Gaps 2;
QY 1562 TGTAGAATAGTCCAGATCTCAGTCCCTGATACATTTATATGCTCTCCATACG 1621
Db 572 TGTAGAATAGTCCAGATCTCAGTCCCTGATACATTTATATGCTCTCCATACG 513
QY 1622 CTATATCAGCAGGAAAGTAGAGTAGGGACATACAGTCTCTTTGTTGCCACAAA 1681
Db 512 CTATATCAGCAGGAAAGTAGAGTAGGGACATACAGTCTCTTTGTTGCCACGAA 453
QY 1682 AATTTTCAGATAACAGCTGGGAAGTCATGATTTGGGTGAGAACTTTGGGGATGTAAGAAA 1741
Db 1741 AATTTTCAGATAACAGCTGGGAAGTCATGATTTGGGTGAGAACTTTGGGGATGTAAGAAA 393
QY 1742 CATTTCTTACAAAAGATCCACCTGCTGCTCCACCGGCGATCGAATAAAGTACA 1801
Db 392 CATTTCTTACAAAAGATCCACCTGCTGCTCCACCGGCGATCGAATAAAGTACA 333
QY 1802 GATTCCTTTTGGGCTGAGCATGTCTAGTATTAACCTTTGCTCTGTTAGGGAAGTGTGG 1861
Db 332 GATTCCTTTTGGGCTGAGCATGTCTAGTATTAACCTTTGCTCTGTTAGGGAAGTGTGG 273
QY 1862 CCATAGATTAGGGTGTAGTTGACAAAACCTTCATCTGGATGTAGTCCAGAAAGTCCCCAC 1921
Db 272 CCATAGATTAGGGTGTAGTTGACAAAACCTTCATCTGGATGTAGTCCAGAAAGTCCCCAC 213
QY 1922 TGCAGGTTAAAGGACACTGGACTCTGCACCTCAGGCACCTAGAGTCTCGCAAGTCTCGGGA 1981
Db 212 TGCAGGTTAAAGGACACTGGACTCTGCACCTCAGGCACCTAGAGTCTCGCAAGTCTCGGGA 153
QY 1982 ACCTGCATTTAAATAAATAATGCATCTATTAATATGTTTCATATCATGTGGCAAAATGGA 2041
Db 152 ACCTGCATTTAAATAAATAATGCATCTATTAATATGTTTCATATCATGTGGCAAAATGGA 94
QY 2042 TAAATTTTGTAGTAACCTTTTAAATTCAGTTGCTGGAATATGGACACATGACCTGGGA 2101
Db 93 TAAATTTTGTAGTAACCTTTTAAATTCAGTTGCTGGAATATGGACACATGACCTGGGA 35
QY 2102 AAATCGTGAATAAATAATGTAAT 2123
Db 34 ATCGTGAATAAATAATGTAATAAT 13

RESULT 6
BG202130/c
LOCUS BG202130 880 bp mRNA linear EST 21-APR-2001
DEFINITION RST21484 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG202130
VERSION BG202130.1 GI:13723817
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 880)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
Offenbacher,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
PUBMED 11329013
COMMENT Contact: Scott J. Cain
Atherys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atherys.com
High quality sequence stop: 384.
FEATURES
source
1..880
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Atherys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression'
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
ORIGIN
```

ORIGIN									
Query Match		16.1%; Score 498; DB 4; Length 880;							
Best Local Similarity		90.1%; Pred. No. 4e-95;							
Matches 589; Conservative		0; Mismatches 59; Indels 6; Gaps 5;							
QY	616	TGCTTTGATGCGGACCGATTTC	CAAGTTTTCAGAGTTTTCAGGAGCAGTCAAGGAAAGGGAAGTTA	675					
DB	825	TGTTGTGTGATGGGACCGATTTC	CAGTTCTGGACCCAGTCAAGGAAAA-GGAAGTAA	767					
QY	676	AGAGCTTTATCCATTTGTTT	TGAGTTGGGACAGAGAAAAATTTGGATTTC	735					
DB	766	AAAGCTTTACCAATTTGGTTT	TGAGTTGGGCCAGAGAAAAATTTGAATTC	708					
QY	736	ATTGATGGAGTCCAAAGTGT	TAGTCTGTTTGGGAGCAGGCTGCTTTAGATCCATTCATA	795					
DB	707	-ATTGAGGAGTCCAAAGTGT	TAGTCTGTTT-GGGAGCAGTGGCTTTTAAACCAATTC	650					
QY	796	CTGTGGAACCTTCTGCGAA	ACTCAACTGACAGATATTACATTTACATGGCTCATTCACA	855					
DB	649	TGTGTAACCTTTGCGCAAA	ACTCAATTCGCAAGTATTCATTTACAATGGCTCATTCACA	590					
QY	856	TCTCT-CCCTGCACAGACA	CAGTTGACTGGATTGTTTTTAAAGATACAGTTAGCATCTC	914					
DB	589	TCTCTCCCGCGCAAGACC	CAGTTGACTGGATTG-TTTTAAAGATACAGTTAGCATCTC	531					
QY	915	TGAAGCCAGTTGGCTGTTT	TTTTGTGAAGTTCTTACAATGCAACAACTCGTTATGTCAT	974					
DB	530	TGAAGCCAGTTGGCTGTTT	TTTTGTGAAGTTCTTACAATGCAACAACTCGTTATGTCAT	471					
QY	975	GCTGATGGACTACTTACA	AAAAAATTTTCGAGAGCAACAGTACAAGTTCTCTAGACAGGT	1034					
DB	470	GCTGATGGACTACTTACA	AAAAAATTTTCGAGAGCAACAGTACAAGTTCTCTAGACAGGT	411					
QY	1035	GTTTTCCTCATACACTG	GGAAGGAGAGATTTCATGAAGCAGTTGTAGTTTCAGAACCCAGA	1094					
DB	410	GTTTTCCTCATACACTG	TAAAGGAGAGATTTCATGAAGCAGTTGTAGTTTCAGAACCCAGA	351					
QY	1095	AAATGTTTCAGGCTGACC	CAGAGAAATTATACAGCCTTCTTGTTACATGGGAAAGACCTCG	1154					
DB	350	AAATGTTTCAGGCTGACC	CAGAGAAATTATACAGCCTTCTTGTTACATGGGAAAGACCTCG	291					
QY	1155	AGTCGTTTATGATACCA	TGATTTGAGAGTTTTCGAGTTTACAGCAGTTGGATGGAGA	1214					
DB	290	AGTCGTTTATGATACCA	TGATTTGAGAGTTTTCGAGTTTACAGCAGTTGGATGGAGA	231					
QY	1215	GGACCAACCAAGCATGA	ATTTTTCAGAGATGGCTATCAAGACTTGGTAACTAT	1268					
DB	230	GGACCAACCAAGCATGA	ATTTTTCAGAGATGGCTATCAAGACTTGGTAACTAT	177					
RESULT 7									
CN383698		701 bp mRNA linear EST 01-OCT-2004							
LOCUS		4117208 BARC_3GAL chicken mixed tissue Gallus gallus cdna clone							
DEFINITION		3GAL_9B14 5', mRNA sequence.							
ACCESSION		CN383698							
VERSION		CN383698.1 GI:53550459							
KEYWORDS		EST.							
SOURCE		Gallus gallus (chicken)							
ORGANISM		Gallus gallus							
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.							
AUTHORS		1 (bases 1 to 701)							
TITLE		EvoC-Clover, C.M., Ashwell, C.M., McMurtry, J.P., Lillehoj, H.S., Matukumalli, L.K. and Van Tassel, C.P.							
JOURNAL		Characterization of expressed sequence tags generated from multiple chicken tissues							
COMMENT		Unpublished (2004)							
		Contact: Christina M. Clover							
		Growth Biology Laboratory							
		Animal and Natural Resources Institute							

Bldg. 200 Rm215 BARC-East, Beltsville, MD 20705, USA									
Tel: 3015048224									
Fax: 3015048623									
Email: chris@nri.barc.usda.gov									
Single pass sequencing. Bases called and trimmed with phred									
0.000925 using options -trim alt -trim fastavector identified by									
cross_match using options -minmatch 12 -minscore 12									
Plate: 9 row: B column: 14									
Seq primer: CCCAGTCACGACGTTGTAACAG									
High quality sequence stop: 701.									
FEATURES									
source									
1..701									
Location/Qualifiers									
/organism="Gallus gallus"									
/mol_type="mRNA"									
/strain="Leghorn and broiler"									
/db_xref="taxon:9031"									
/clone="3GAL 9B14"									
/lab_host="DH5alpha"									
/clone_lib="BARC 3GAL chicken mixed tissue"									
/note="vector: pBluescript, SK-, Stratagene; Site_1: NotI;									
Site_2: EcoRI; Normalized library of pooled RNA isolated									
from whole brain, ultimobronchial gland, parathyroid									
gland, cecal tonsil and primordial germ cells Multiple"									
ORIGIN									
Query Match 16.0%; Score 494.8; DB 7; Length 701;									
Best Local Similarity 81.8%; Pred. No. 1.9e-94;									
Matches 571; Conservative 0; Mismatches 127; Indels 0; Gaps 0;									
Qy	568	GAGCATAGTTTAGAAGGACAA	AAATTTCCACTTCAGATGCAAAATCTACTGCTTTGATGCG	627					
Db	4	GAACATAGCTTGGAGAGCA	AAATTTCTCTTGGATGCAAAATCTACTGCTATGATGGA	63					
Qy	628	GACCGATTTTCAAGTTT	TGAGGAGCAGTCAAAGGAAAGGAGTTAAGAGCTTTATCC	687					
Db	64	GATCTGTTTACAGACTT	CAAGAGGCAATTAAGAGAAATGAAAGCTTTATCA	123					
Qy	688	ATTTTGTGAGGTTGG	CACAGAGAAATTTGGAATTCAAAGCGATTATGATGGAGTC	747					
Db	124	GTTTTGTGAGATTGG	AGTAGAAGATAATCCGGATTATATTCCAATCATTTAAGGGAGTA	183					
Qy	748	GAAAGTGTAGTCCG	TTTTTGGGAAGCAGCTGCTTTAGATCCATTCATATGTTGAACCTT	807					
Db	184	GATAGTGTAGTCCG	TTTTTGGGAACACAGCTGCTTAGAATTTGTTGCTGACCTT	243					
Qy	808	CTGCCAAATCAACT	GAAGTATTAATTTCAATGGCTCATTTGACATCTCTCCCTGCG	867					
Db	244	TTACCGAATGCA	CGGACAAATATTACATTTACAAATGGGCTTTTATCAGCTCTCTCCCTGC	303					
Qy	868	ACAGACAGTTGACT	CGGATTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG	927					
Db	304	TCAGAAACAGTT	GGAATTTGTTTCAAGATATCTATTAGTATTTCTGAGCAACAGTTA	363					
Qy	928	GCTGTTTTTGTGA	AGTTCTTTTACAAATGCAAAATCTGGTTATGTCATGCTGATGACTAC	987					
Db	364	GCATATTTCTG	TGAAGTCCCTTACAAATGCAACAGCTCTGGCTATGATGCTCATGACTAT	423					
Qy	988	TTACAAACAAATTT	TCGAGAGCAACAGTACAAAGTTCTCTAGACAGGTTGTTTTCTCATAC	1047					
Db	424	CTGCAAAACAACT	TTCCGAGAGCAACGATATAAGTTCTCTGCGCAAGTGTGTTTTCTCTTAC	483					
Qy	1048	ACTGGAAGGAGAG	ATTTCATGAGCAGTTTGTAGTTTCAGAACCCAGAAATGTTTCAGCT	1107					
Db	484	ACTGGACAGGA	AGAAATTTTCATGAAGCAGTTTTCAGAACCTCAGAAATGTCCAATCT	543					
Qy	1108	GACCCAGAGAA	TATACAGCTTTCTTGTGTACATGGGAAACCTCGAGTGTGTTTATGAT	1167					
Db	544	GATCCAAAGA	ATATATACAGCTTCTGTTACGTTGGGAAAGACCTCGGTTGTGTATGAT	603					
Qy	1168	ACCATGATGAG	AAAGTTTTCAGGTTTGTATACAGAGTTTGGATGGAGGACCAACCCAG	1227					
Db	604	ACCATGATGAG	AAAGTTTTCAGGTTTGTATACAGAGTTTGGATGGAGGACCAACCCAG	663					

QY 1228 CATGAATTTTTCACAGATGCTATCAAGACTTGGTAAC 1265
 |||||
 Db 664 CATGAGTTTCTGACAGATGGTATCAGGACTTGGGAGC 701
 |||||

RESULT 8

CA879683 585 bp mRNA linear EST 20-DEC-2002
 LOCUS K0976A09-5N NIA Mouse Neural Stem Cell (Undifferentiated) cDNA
 DEFINITION Library (Long) Mus musculus cDNA clone NIA:K0976A09 IMAGE:30091496
 5', mRNA sequence.
 ACCESSION CA879683
 VERSION
 KEYWORDS
 SOURCE CA879683.1 GI:27331232
 ORGANISM Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 585)
 AUTHORS Piao, Y., Dudekula, D.B., Qian, Y., Martin, P.R., Alba, K., Vescovi, A.L.
 and Ko, M.S.H.
 TITLE Systematic Analyses of NIA Mouse Neural Stem Cell
 (Undifferentiated) cDNA Library (Long)
 JOURNAL Unpublished (2002)
 COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: K0976 row: A column: 09
 Seq primer: -21M13 Reverse
 High quality sequence stop: 585
 POLYA=No.

FEATURES

Location/Qualifiers
 1..585
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CD1"
 /db_xref="niaEST:K0976A09-5N"
 /db_xref="taxon:10090"
 /clone="NIA:K0976A09 IMAGE:30091496"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated)
 cDNA Library (Long)"
 /note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
 a long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
 obtained from Dr. Angelo L. Vescovi (Institute for Stem
 Cell Research, Italy). Double-stranded cDNAs were
 synthesized with an Oligo(dT) primer [Invitrogen:
 5'-pCAGTAGTCTAGTCGCGCGCCCTTTT-3'] from
 2.0 Microgram of total RNA, treated with T4 DNA
 polymerase, and purified by ethanol-precipitation. The
 cDNAs were ligated to Lene-linker LL-SalI, purified by
 phenol/chloroform, and separated from free linkers by
 Centricion 100. Then, the cDNAs were amplified by
 long-range high fidelity PCR using Ex Taq polymerase
 (Takara) with a primer Sal4-S. The products were purified
 by phenol/chloroform and Centricon 100. The cDNAs were
 digested with SalI and NotI enzymes and cloned into
 SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.
 coli host was transformed with the ligation mixture by the
 standard chemical method. The average insert size is about
 3.8 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 15.7%; Score 484.2; DB 6; Length 585;
 Best Local Similarity 89.2%; Pred. No. 3.3e-92;
 Matches 522; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 355 CTTACACAAGTAATGTGAATCTTAAGAAACTTAAATTTTCAGGGTTGGGATAAAACATCA 414
 |||||
 Db 1 CTTACACAAGTCAATGTGAATCTTAAGAACTGAATTTTCAGGGTTGGGAAAGGCTCC 60
 |||||
 QY 415 TTGGAAAAACATTCATTCAATCAACAACCTGGGAAAAACAGTGGAAATTAATCTCACTAATGAC 474
 |||||
 Db 61 TTGGAAAAACATTCATTCAATCAACAACCTGGGAAAAACAGTGGAAATTAATCTCACTAATGAC 120
 |||||
 QY 475 TACCGTGTTCAGGGAGGAGTTTCAGAAATGTGTTTAAAGCAGCAAGATACTTTTTCAC 534
 |||||
 Db 121 TACTATCTCAGTGGGAGGACTTTTCAGAAAAAGGCTTCAAGGCGCAAGCAAGATATCTTTCCAC 180
 |||||
 QY 535 TGGGAAAAATGCAATATATGTCTCATCTGATGATCAGAGCATAGTTTAAAGAGGACAAAAATTT 594
 |||||
 Db 181 TGGGAAAAATGCAATGTGTCTCATCTGAGGATCGGAACTAGCTTAGAAGGACAGAGTTTC 240
 |||||
 QY 595 CCACCTTGAGATGCAAACTACTGTCTTTCATGATCGGACCGATTTTCAAGTTTTCAGGAAAGCA 654
 |||||
 Db 241 CCACCTTGAGATGCAAGTCTACTGTCTTTCATGATCGGACAGATTTTCCAGTTTTCAGGAAAGCA 300
 |||||
 QY 655 GTCAAAGGAAAAAGGAGTAAAGAGCTTTATCCATTTTGTGAGTTGGGACAGAA 714
 |||||
 Db 301 GTTAAAGGAAAAAGGAGATTAAGGCTTTATCCATTTTATTTGAGTTGGAGTTTGAAGAA 360
 |||||
 QY 715 AATTGGATTTCAAAGCGATTAATTGATGAGTTCGAAAGTGTAGTCTGTTTGGGAAAGCAG 774
 |||||
 Db 361 AATTGGATTTCAAAGCGATTAATTGATGAGTTCGAAAGTGTAGTCTGTTTGGGAAAGCAG 420
 |||||
 QY 775 GCTGCTTTAGATCCATTCATCTGTTGAACCTTTCTGCCAAACTCAACTGACAGATTTTAC 834
 |||||
 Db 421 GCTGCTTTAGATCCATTCATCTGTTGAACCTTTCTGCCAAACTCAACTGACAGATTTTAC 480
 |||||
 QY 835 ATTACATATGCTCATTCAGATCTCTCCCTCGCAGACAGATTCAGTGGATTTGTTT 894
 |||||
 Db 481 ATTACATATGCTCATTCAGATCTCTCCCTCGCAGACAGATTCAGTGGATTTGTTT 540
 |||||
 QY 895 AAAGATACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTGT 939
 |||||
 Db 541 AAGATACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTGT 585
 |||||

CK781524 718 bp mRNA linear EST 23-FEB-2004
 LOCUS UI-M-GV0-cly-h-20-0-UI.r1 NIH_BMAP_GV0 Mus musculus cDNA clone
 DEFINITION IMAGE:30622171 5', mRNA sequence.

ACCESSION CK781524

VERSION CK781524.1 GI:42747202

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 718)

AUTHORS NIH-MSC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cga@bbs-r@mail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

http://genome.uiowa.edu/distribution/mousefi.html

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..718

/organism="Mus musculus"

/mol_type="mRNA"

FEATURES

source

/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30622171"
/tissue_type="whole brain"
/dev_stage="1.5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP GVO"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAAGTGAAT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN
Query Match 15.6%; Score 481.6; DB 7; Length 718;
Best Local Similarity 88.3%; Pred. No. 1.2e-91;
Matches 523; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 677 GAGCTTTTATCCATTTGTTGAGGTTGGGACAGAGAAATTTGGATTTCAAAGCGATTA 736
DB 1 GGGCTTTATCCATTTATTTGAGGTTGGAGTTGAAGAAATTTGGATTTCAAAGCGATTA 60
QY 737 TTGATGGAGTGGAAAGTGTATGCTTTTGGAGAGCAGGCTGCTTTAGATCATTTATCAT 796
DB 61 TTGATGGAACTGAGAGTGTATGCTTTTGGAAAGCAGGCTGCTTTAGATCATTTGCTCT 120
QY 797 TGTGAACCTTCTGCCAACTCACTGACAGTATTACATTTACATGCTCATTCAGAT 856
DB 121 TGCAGAACCTCTGCGCAACTCCACTGACAGTATTACATTTACATGAGTCAATTCAGAT 180
QY 857 CTCCTCCCTGCACACACAGTGTGACTGGATTTGTTTTAAAGATACAGTTAGCATCTCTG 916
DB 181 CCCCTCCCTGCACACACACCGTGGAAATGATTTGTTTTAAAGATACAGTTAGCATCTCTG 240
QY 917 AAAACCGAGTGGCTGTTTTTGTGAAGTCTTACAAATGCAACAATCTGGTATGTCATGCG 976
DB 241 AAAACCGAGTGGCTGTTTTTGTGAAGTCTTACAAATGCAACAATCTGGTATGTCATGCT 300
QY 977 TGATGGAATCTTACAAACAATTTCCGAGAGCAAGTCAAGTTCTCTAGACAGGTGT 1036
DB 301 TGATGGAATCTTACAAACAATTTCCGAGAGCAAGTCAAGTTCTCTAGACAGGTGT 360
QY 1037 TTTCTCTATACACTGGAAGGAGAGATTCATGAAGCAGTTTGTAGTTTCAGAACACAGAAA 1096
DB 361 TTTCTCTATACTTGGAAAGGAGAGATCCAGAGTAGTGTAGTTTCAGAACACAGAAA 420
QY 1097 ATGTTACGGCTGACCCAGAGAAATTAACAGGCTTTCTTTGTATCATGGAAGAGCTTCGAG 1156
DB 421 ATGTCGAAGCTGACCCCTGAGAATTAACAGGCTTTCTGGTCAATGGAAGAGCTTCGAG 480
QY 1157 TCGTTTATGATACCATGATTCGAGAGTTTTCGAGTTTTCGACAGGTTGATGAGAGG 1216
DB 481 TCGTTTATGAGCCATGATTCGAGAGTTTTCGAGTTTTCGACAGGTTGATGAGAGG 540
QY 1217 ACCAAACCAAGCATGAAATTTTGAAGAGTGGCTATCAAGACTTTGGTAACTAT 1268
DB 541 ACCAAGCCAGCATGAGTTCTTAACAGATGGCTATCAGGACTTTGGTGCCAT 592

RESULT 10
AL706826 695 bp mRNA linear EST 04-SEP-2003
LOCUS AL706826
DEFINITION DKFZp686J1044_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 695)
Lauber, J., Bahr, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fob, G., Han, M. and Wiemann, S.
EST (Lauber, J., Bahr, A., Mewes, H.W., Weil, B., Amid, C., et al.)
Unpublished (2003)
Contact: MIPS

COMMENT

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No 5' sequence available.

This clone (DKFZp686J1044) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..695
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686J1044"
/dev_stage="adult"
/lab_host="DH10B"
/note="vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN

Query Match 15.4%; Score 475.4; DB 1; Length 695;
Best Local Similarity 99.8%; Pred. No. 2.5e-90;
Matches 476; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CACACATACGACGACGACGATCTCACTTCGATCTATACACTGGAGGATTTAAACAAACAAA 60
DB 219 CACACATACGACGACGACGATCTCACTTCGATCTATACACTGGAGGATTTAAACAAACAAA 278
QY 61 CAAAAAACAATTTCTTCGCTCCCTCCCTCTCCACTCTGAGAAGCAGAGGAGCGCA 120
DB 279 CAAAAAACAATTTCTTCGCTCCCTCCCTCTCCACTCTGAGAAGCAGAGGAGCGCA 338
QY 121 CGGCGAGGGCGGACGACCGTCTGGAATCGGAATCTTAAAGCGTTTCCTCGCTTGCATT 180
DB 339 CGGCGAGGGCGGACGACCGTCTGGAATCGGAATCGGAATCGGAATCGGAATCGGAAT 398
QY 181 CAGCTCTCTGTGTTTGGCGCTGATTTGGCTAAATGATACAGCAACAGAGAAA 240
DB 399 CAGCTCTCTGTGTTTGGCGCTGATTTGGCTAAATGATACAGCAACAGAGAAA 458
QY 241 CTTGTTGAAGAGATTTGGCTGCTCTATACAGGAGCAGTGAATCAAAAAAATTTGGGGAAG 300
DB 459 CTTGTTGAAGAGATTTGGCTGCTCTATACAGGAGCAGTGAATCAAAAAAATTTGGGGAAG 518
QY 301 AAATATCAACATGTAATAGCCCAAAACAATCTCTATCAATATGATGAGATCTTACA 360
DB 519 AAATATCAACATGTAATAGCCCAAAACAATCTCTATCAATATGATGAGATCTTACA 578
QY 361 CAAGTAAATGGAATCTTAAAGAAATTTAAATTTTCAAGGGTTGGGATAAAAATTTGGA 420
DB 579 CAAGTAAATGGAATCTTAAAGAAATTTAAATTTTCAAGGGTTGGGATAAAAATTTGGA 638
QY 421 AACACATTCATTAACACATCGGGAACAGTGGAAATTAATCTCACTTAATGACTAC 477
DB 639 AACACATTCATTAACACATCGGGAACAGTGGAAATTAATCTCACTTAATGACTAC 695

RESULT 11
 CA874874
 LOCUS
 DEFINITION
 CA874874
 CA874874
 CA874874.1 GI:27326423
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 664)
 Piao Y., Dudekula D.B., Qian Y., Martin P.R., Aiba K., Vescovi A.L.
 and Ko M.S.H.
 Systematic Analyses of NIA Mouse Neural Stem Cell
 (Undifferentiated) cDNA Library (Long)
 Unpublished (2002)
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@gsun.grc.nia.nih.gov
 Plate: K0935 row: B column: 06
 Seq primer: -21M13 Reverse
 High quality sequence stop: 664
 POLYA=No.

FEATURES

Location/Qualifiers
 1..664
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Cbl"
 /db_xref="taxon:10090"
 /db_xref="taxon:10090"
 /clone="NIA:K0935R06 IMAGE:30087605"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated)
 cDNA Library (Long)"
 /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
 a long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
 obtained from Dr. Angelo L. Vescovi (Institute for Stem
 Cell Research, Italy). Double-stranded cDNAs were
 synthesized with an oligo(dT) primer [Invitrogen:
 5'-TGACTAGTCTAGATCGGAGCGCGCCCTTTT-3'] from
 2.0 microgram of total RNA, treated with T4 DNA
 polymerase, and purified by ethanol-precipitation. The
 cDNAs were ligated to Lone-linker LL-Sal4, purified by
 phenol/chloroform, and separated from free linkers by
 Centricon 100. Then, the cDNAs were amplified by
 long-range high fidelity PCR using Ex Taq polymerase
 (Takara) with a primer Sal4-S. The products were purified
 by phenol/chloroform and Centricon 100. The cDNAs were
 digested with SalI and NotI enzymes and cloned into
 SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.
 coli host was transformed with the ligation mixture by the
 standard chemical method. The average insert size is about
 3.8 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 15.3%; Score 473.2; DB 6; Length 664;
 Best Local Similarity 85.0%; Pred. No. 7.3e-90;
 Matches 529; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
 73 TTCCTTCGCTCCCTCTCCACTCTGAGAGCAGAGGCGCGCGGCGGCGCC 132
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 43 TGCCTCGGTGTCCACCCCTCTGTACCCCTGGCGTGGCGGGCGC 102
 Qy 133 GCAGACCGTCTGGAAATGCGAATCCTTAAAGCGTTTCTCGCTTCGCTTCATTCAGCTCTCTGT 192
 Db 103 CGGACCGGCTGGAGATCGAATCTCTGAGAGCTTCTCGCGTGGTTCAGCTCTGTGC 162
 Qy 193 GTTTCGCGCTGGATTGGCTTAATGGATCTACAGACAACAGAGAAACTTTGTTGAAGAG 252
 Db 163 CTGTGTGCGCTGGAGTGGCTTATGGATCTACAGACAACAGAGAAACTTTGTTGAAGAG 222
 Qy 253 ATTGGCTGCTCTATACAGGAGCTGATCAAAAATAATTTGGGAAAGAAATATCCAACA 312
 Db 223 ATTGGCTGCTCTACAGAGGAGCTAAATCAAAAATAATTTGGGAAAGAAATATCCAATA 282
 Qy 313 TGTATAGCCCCAAACAACTCTCTATCAATATGATGAAGATCTTACACAAGTAAATCTG 372
 Db 283 TGTATAGCCCCAAAGAGCTCTCTATTAATATGATGAAGATCTTACACAAGTCAATGTG 342
 Qy 373 AATCTTAAGAACTTAAATTTTCAGGCTTGGGATAAAACATCATTTGGAAAAACACATTCATT 432
 Db 343 AATCTTAAGAACTTAAATTTTCAGGCTTGGGATAAAAGCGTCTTGGAAAAACACGTTCA 402
 Qy 433 CATTAACACTGGGAAAAACAGTGGAAATTAATCTCACTAATGATACCTGTCAGCGGAGGA 492
 Db 403 CACAACACTGGGAAAAACAGTGGAAATTAATCTCACTAATGATACCTGTCAGTGGAGGA 462
 Qy 493 GTTTTCAGAAATGGTGTTTTAAAGCAAGCAAGTAATCTTCACTGGGAAAAATGCAATATG 552
 Db 463 CTTTTCAGAAAGGTCTTCAAGGCAAGCAAGTAATCTTCACTGGGAAAAATGCAATATG 522
 Qy 553 TCATCTGATGATCAGAGCATAGTTTAAAGAGCAAGAAATTTCCACTTGAGATGCAAAATC 612
 Db 523 TCATCTGAGGATCGGAACATAGCTTAGAAGGACAGAAGTTCCCACTGGAGATCAAGTC 582
 Qy 613 TACTGCTTTGATGCGGACCGGATTTTCAAGTTTTCAGGAGAGCAGTCAAGAGAAAGGGAAG 672
 Db 583 TACTGCTTTGATGCGGACAGATTTTCCAGTTTTCAGGAGAGCAGTCAAGAGAAAGGGAAGA 642
 Qy 673 TTAAGAGCTTTATCATTTTGT 694
 Db 643 TTAAGGCTTTATCATTTTAT 664

RESULT 12
 CD803608
 LOCUS
 DEFINITION
 UI-M-GVO-chv-f-19-0-UI.r1 NIH BMAP_GVO Mus musculus cDNA clone
 IMAGE:30544938 5', mRNA sequence.
 CD803608
 ACCESSION
 CD803608.1 GI:32462434
 VERSION
 EST.
 SOURCE
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 780)
 NIH-MSC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousefl.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pYX-5.
 Location/Qualifiers
 1..780
 /organism="Mus musculus"

FEATURES

source

/mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30544938"
 /issue_type="whole brain"
 /tissue_type="1, 5, and 15 days newborn"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_GVO"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAAGTGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match	14.7%;	Score 453.6;	DB 6;	Length 780;
Best Local Similarity	88.5%;	Pred. No. 1.1e-85;	Mismatches 64;	Indels 0;
Matches 492;	Conservative 0;			

Qy	713	AAAATTTGGATTTC	CAAGCGATTATTGATGAGT	CGAAGTGTAGTGTTCGGGAGC	772
Db	10	AAATTTGGATTTC	CAAGCGATTATTGATGAGT	CGAAGTGTAGTGTTCGGGAGC	69
Qy	773	AGGCTGCTTTAGAT	TCATCATCTGTTGAACCTTCTGCCAACTCACTGACAAAGTATT	832	
Db	70	AGGCTGCTTTAGAT	TCATCATCTGTTGAACCTTCTGCCAACTCACTGACAAAGTATT	129	
Qy	833	ACATTTACAATGGCT	CATTGACATCTCTCCCTGCACAGACACAGTTGACCTGGATGTTT	892	
Db	130	ACATTTACAATGGCT	CATTGACATCTCTCCCTGCACAGACACAGTTGACCTGGATGTTT	189	
Qy	893	TTAAGATACAGT	TAGCATCTCTGAAAGCCAGTTGGCTGTTTTTGTGAAGTCTTACAA	952	
Db	190	TTAAGATACAGT	TAGCATCTCTGAAAGCCAGTTGGCTGTTTTTGTGAAGTCTTACAA	249	
Qy	953	TGCAACAATCTGGT	TATGTCATGCTGATGGAATCTTACAAAAACAATTTTCGAGAGCAAC	1012	
Db	250	TGCAACAATCTGGT	TATGTCATGCTGATGGAATCTTACAAAAACAATTTTCGAGAGCAAC	309	
Qy	1013	AGTACAGTCTCT	TAGACAGCTGTTTCTCTCATACCTGGAAGGAGAGATTCATGAAG	1072	
Db	310	AGTACAGTCTCT	TAGACAGCTGTTTCTCTCATACCTGGAAGGAGAGATTCATGAAG	369	
Qy	1073	CAGTTTGTAGTTC	AGAACAGAAATGTTTCAGGCTGACCCAGAGAAATATACAGCCTTC	1132	
Db	370	TAGTGTGTAGTTC	AGAACAGAAATGTTTCAGGCTGACCCAGAGAAATATACAGCCTTC	429	
Qy	1133	TGTTTACATGGGAA	AGACCTCGAGTCGTTTATGATACCATGATTGAGAGTGTTCAGTGT	1192	
Db	430	TGTTTACATGGGAA	AGACCTCGAGTCGTTTATGATACCATGATTGAGAGTGTTCAGTGT	489	
Qy	1193	TGTACACAGCT	TGGATGGAGGACCAACACAGCATGATGATTTTGTACAGATGCTATC	1252	
Db	490	TGTACACAGCT	TGGATGGAGGACCAACACAGCATGATGATTTTGTACAGATGCTATC	549	
Qy	1253	AAGACTTGGTAACT	AT 1268		
Db	550	AGGACTTGGTGGCAT	565		

RESULT 13
 CA870557
 LOCUS

DEFINITION K0903A03-5N NIA Mouse Neural Stem Cell (Undifferentiated) cDNA Library (Long) Mus musculus cDNA clone NIA:K0903A03 IMAGE:30084482

5', mRNA sequence.
 CA870557
 EST.
 CA870557.1 GI:27322106

Mus musculus (house mouse)
 Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (Bases 1 to 522)

AUTHORS Piao Y., Dudekula D.B., Qian Y., Martin P.R., Aiba K., Vescovi A.L. and Ko M.S.H.

TITLE Systematic Analyses of NIA Mouse Neural Stem Cell (Undifferentiated) cDNA Library (Long)

JOURNAL Unpublished (2002)

COMMENT Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@igsun.grc.nia.nih.gov
 Plate: K0903 row: A column: 03
 Seq primer: -21M13 Reverse

High quality sequence stop: 522
 POLYA=No.

FEATURES

Location/Qualifiers	source
1..522	/organism="Mus musculus"
/mol_type="mRNA"	
/strain="CD1"	
/db_xref="niaEST:K0903A03-5N"	
/db_xref="taxon:10090"	
/clone="NIA:K0903A03 IMAGE:30084482"	
/lab_host="Adult"	
/dev_stage="Adult"	
/clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated) cDNA Library (Long)"	
/note="vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Angelo L. Vescovi (Institute for Stem Cell Research, Italy). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTTCGATCGGAGCGCCGCTTTTCTTTT-3'] from 2.0 Microgram of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker Lu-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.8 kb. The library was constructed by Yulan Piao."	

ORIGIN

Query Match	13.9%;	Score 430.8;	DB 6;	Length 522;
Best Local Similarity	89.1%;	Pred. No. 7.4e-81;	Mismatches 57;	Indels 0;
Matches 465;	Conservative 0;			

Qy	564	ATCAGAGCATAGTTT	TAGAAGGACAAAAATTTCCACTTTCAGATGCAAAATCTACTGCTTTGA	623
Db	1	ATCGGACATAGCTT	TAGAAGGACAGAGTTCCTCCACTGAGATGCAAGTCTACTGCTTTGA	60
Qy	624	TGCGGACCGATTTT	CAAGTTTGTAGGAGCAGTCAAGGAAAGGAGGAGTTTAAGAGCTTT	683
Db	61	TGCGGACAGATTTT	CCAGTTTGTAGGAGCAGTCAAGGAAAGGAGGAGTTTAAGAGCTTT	120


```
QY 684 ATCCATTTTGTGAGGTTGGGACAGACGAAATTTGGATTTCAAGCGATTATTGATGG 743
Db 121 ATCCATTTTGTGAGGTTGGGAGTTGGAATTTGAAGAAATTTGGATTACAAAGCCATTATTGATGG 180
QY 744 AGTGGAAAGTGTAGTCGTTTGGGAGCAGGCTGCTTTAGATCCATTCATCTGTTGAA 803
Db 181 AACTGAGAGTGTAGTCGTTTGGGAGCAGGCTGCTTTAGATCCATTCATCTGTTGAGAA 240
QY 804 CTTCTGCGCAACTCAACTGACAAGTATTACATTTTACAATGGCTCAATGACATCTCCCTCC 863
Db 241 CTTCTGCGCAACTCAACTGACAAGTATTACATTTTACAATGGCTCAATGACATCTCCCTCC 300
QY 864 CTGCACAGACAGATTGACTGGAATGTTTAAAGATACAGTTAGCATCTCTGAAAGCCA 923
Db 301 CTGCACAGACAGCGTGAATGGAATGTTTAAAGATACAGTTAGCATCTCTGAAAGCCA 360
QY 924 GTTGGCTGTTTTGTGAAAGTCTTACAATGCAACATCTGGTTATGTCATGCTGATGGA 983
Db 361 GCTGGCTGTATTTGTGAAAGTCTTACAATGCAACAGTCTGGGTATGTCATGTTGATGGA 420
QY 984 CTACTTTACAAAACAATTTTCGAGAGCAACAGTACAGTTCTTACAGAGGTTTTCCTC 1043
Db 421 TTACTTTACAAAACAATTTTCGAGAGCAACAGTACAGTTCTTACAGAGGTTTTCCTC 480
QY 1044 ATACACTGGAAAGGAGAGATTTCATGAAGCAGTTTGTAGTTC 1085
Db 481 ATATACTGGAAAGGAGAGATCCAGAGTAGTGTGTAGTTC 522

RESULT 14
BX642570 629 bp mRNA linear EST 04-SEP-2003
LOCUS DKFZp686L2091_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DEFINITION DKFZp686L2091 5', mRNA sequence.
ACCESSION BX642570
VERSION BX642570.1 GI:34476903
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 629)
AUTHORS Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
TITLE EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
Wellenreuther,R., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No si sequence available.
This clone (DKFZp686L2091) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
source 1..629
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686L2091"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"
```

```
ORIGIN
Query Match 13.4%; Score 414.6; DB 5; Length 629;
Best Local Similarity 98.8%; Pred. No. 2.1e-77;
Matches 428; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 1 CACACATACGACGACGATCTCACTTCGATCTATACACTCGAGGATTAAAAACAACAAA 60
Db 198 CACACATACGACGACGATCTCACTTCGATCTATACACTCGAGGATTAAAAACAACAAA 257
QY 61 CAAAAAAACAATTCCTTCGCTCCCTCCCTCTCCACTCTGAGAGGAGGCGCGCA 120
Db 258 CAAAAAAACAATTCCTTCGCTCCCTCCCTCTCTCCACTCTGAGAGGAGGCGCGCA 317
QY 121 CGCGAGGGCGCGAGACCGCTCGGAATGGAATCCCTAAAGCGTTTCTCGCTTCGATT 180
Db 318 CGCGAGGGCGCGAGACCGCTCGGAATGGAATCCCTAAAGCGTTTCTCGCTTCGATT 377
QY 181 CAGCTCCTCTCTGTTTGGCGCTGGATTGGGCTTAATGGATCTACAGACAACAGAGAAAA 240
Db 378 CAGCTCCTCTCTGTTTGGCGCTGGATTGGGCTTAATGGATCTACAGACAACAGAGAAAA 437
QY 241 CTGTTGAAGAGATGCGTCTCTATACAGAGACATCGAATCAAAAAAATTTGGGGAAG 300
Db 438 CTGTTGAAGAGATGCGTCTCTATACAGAGACATCGAATCAAAAAAATTTGGGGAAG 497
QY 301 AAATATCCAACTGTAATAGCCCAAAACAATCTCTATCAATATATGATGAAGATCTTACA 360
Db 498 AAATATCCAACTGTAATAGCCCAAAACAATCTCTATCAATATATGATGAAGATCTTACA 557
QY 361 CAAGTAAATGTGAATCTTAAAGAACTTAAATTCAGGGTTGGGATAAAACATCATTTGAA 420
Db 558 CAAGTAAATGTGAATCTTAAAGAACTTAAATTCAGGGTTGGGATAAAACATCATTTGAA 616
QY 421 AACACATTCATTC 433
Db 617 AACACATTCATTC 629

RESULT 15
CA872941 498 bp mRNA linear EST 20-DEC-2002
LOCUS K0921G05-5N NIA Mouse Neural Stem Cell (undifferentiated) cDNA
DEFINITION Library (Long), Mus musculus cDNA clone NIA.K0921G05 IMAGE:30086284
5', mRNA sequence.
ACCESSION CA872941
VERSION CA872941.1 GI:27324490
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 498)
AUTHORS Piao,Y., Dudekula,D.B., Qian,Y., Martin,P.R., Aiba,K., Vescovi,A.L.
and Ko,M.S.H.
TITLE Systematic Analyses of NIA Mouse Neural Stem Cell
(Un differentiated) cDNA Library (Long)
JOURNAL Unpublished (2002)
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: K0921 row: G column: 05
Seq primer: -21M13 Reverse
High quality sequence stop: 498
POLYA=No.
FEATURES
Location/Qualifiers
source 1..498
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD1"
/db_xref="niaEST:K0921G05-5N"
```

Job time : 9636 secs

```
/db_xref=taxon:10090"
/clone=NIA:K092IG05 IMAGE:30086284"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated)
cDNA Library (Long)"
/notes="vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://igun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Dr. Angelo L. Vescovi (Institute for Stem
Cell Research, Italy). Double-stranded cDNAs were
synthesized with an Oligo(dT) primer [Invitrogen:
5'-PGACTAGTCTAGATCGGCGGCCCTTTT-3'] from
2.0 Microgram of total RNA, treated with T4 DNA
polymerase, and purified by ethanol-precipitation. The
cDNAs were ligated to Lone-linker LL-Sal4, purified by
phenol/chloroform, and separated from free linkers by
Centricon 100. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S. The products were purified
by phenol/chloroform and Centricon 100. The cDNAs were
digested with SalI and NotI enzymes and cloned into
SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.
coli host was transformed with the ligation mixture by the
standard chemical method. The average insert size is about
3.8 kb. The library was constructed by Yulan Piao."
```

ORIGIN

Query Match		13.2%;	Score 409;	DB 6;	Length 498;
Best Local Similarity		88.9%;	Pred. NO. 3.2e-76;		
Matches 442;		Conservative 0;	Mismatches 55;	Indels 0;	Gaps 0;
QY	451	GTGGAAATTAATCTC	ACCTAATGACTACCGT	CTCAGCGGAGGAGTTT	CAGAAATCGTGTTT 510
Db	1	GTGGAAATTAATCTC	ACCTAATGACTACTAT	CTCAGTGGGAGCTTT	CAGAAAAGGTTTC 60
QY	511	AAAACAAGCAAGATA	CTTTTTCACCTGGGAAA	ATGCAATATGTCTAT	CTGTGATGATCAGAG 570
Db	61	AAGACAAGCAAGATA	CTTTTTCACCTGGGAAA	ATGCAATATGTCTAT	CTGTGATGATCGAA 120
QY	571	CATAGTTTAGAAGN	CAAAATTTCCACTTC	GAGATGCAAAATCTA	CTGCTTTGATCGGAC 630
Db	121	CATAGTTTAGAAGN	CAAGATTTCCCACTG	GAGATGCAAAATCTA	CTGCTTTGATCGGAC 180
QY	631	CGATTTTCAAGTTT	TGAGGAGCAGTCAA	AGGAAAGGAAAGTT	TAAGAGCTTTATCCATT 690
Db	181	AGATTTTCAAGTTT	TGAGGAGCAGTCAA	AGGAAAGGAAAGTT	TAAGAGCTTTATCCATT 240
QY	691	TTGTTGAGTTGGG	ACAGAAATTTGGATT	TTCAAAGCGATTATT	TGATGGAGTCGAA 750
Db	241	TTATTTGAGTTGG	AGTTTGAAGAAATTT	GGATTACAAAGCCAT	TATTGATGGAACTGAG 300
QY	751	AGTCTTAGTCTGT	TTGGGAAGCAGCTCT	TTAGATCCATTCTA	CTGTTGAACTTCTG 810
Db	301	AGTCTTAGTCTGT	TTGGGAAGCAGCTCT	TTAGATCCATTCTG	TCGTGTCAGAACCTCCTG 360
QY	811	CCAAACTCACTG	ACAAGTATTACATTT	TACATGCTCATTT	GACATCTCTCCCTGCACA 870
Db	361	CCAAACTCACTG	ACAAGTATTACATTT	TACATGCTCATTT	GACATCTCTCCCTGCACA 420
QY	871	GACACAGTTGAT	GGATTTGTTTTTAA	GATACAGTTAGCAT	CTCTGAAAGCCAGTTGGCT 930
Db	421	GACACCGTGAAT	GGATTTGTTTTTAA	GATACAGTTAGCAT	CTCTGAAAGCCAGTTGGCT 480
QY	931	GTTTTGTGTA	AGTTCT 947		
Db	481	GTATTTGTGA	AGTTCT 497		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 11:21:09 ; Search time 1579 Seconds
(without alignments)
11588.297 Million cell updates/sec

Title: US-09-983-000A-1
Perfect score: 3091
Sequence: 1 cacacatgcgcacgcagcgt.....ttatgtagtaagtgatcata 3091

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980a:*
2: Geneseqn1990a:*
3: Geneseqn2000a:*
4: Geneseqn2001a:*
5: Geneseqn2001b:*
6: Geneseqn2002a:*
7: Geneseqn2002b:*
8: Geneseqn2003a:*
9: Geneseqn2003b:*
10: Geneseqn2003c:*
11: Geneseqn2003d:*
12: Geneseqn2004a:*
13: Geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3091	100.0	3091	8	Aad48126 Human pro
2	1263.2	40.9	5367	8	Abx76417 Lung canc
3	1263.2	40.9	5367	8	Abx76218 Lung canc
4	1263.2	40.9	5367	11	Adn39055 Cancer/an
5	1263.2	40.9	5367	11	Adn39726 Cancer/an
6	1263.2	40.9	7941	6	Abi66502 Lung canc
7	1263.2	40.9	7941	8	Abx76421 Lung canc
8	1263.2	40.9	7941	8	Aad48125 Human PTP
9	1263.2	40.9	7941	10	Acc72641 Human pro
10	1263.2	40.9	7941	11	Adn39063 Cancer/an
11	1263.2	40.9	7941	12	Adj75139 Marker ge
12	1263.2	40.9	7944	8	Abx76217 Lung canc
13	1263.2	40.9	7944	8	Abx76416 Lung canc
14	1263.2	40.9	7944	11	Adn39725 Cancer/an
15	1263.2	40.9	7944	11	Adn39053 Cancer/an
16	1263.2	40.9	8058	8	Aad48127 Human pro
17	1263.2	40.9	8292	12	Adq22702 Human sof
18	1261.6	40.8	5343	8	Abx76418 Lung canc
19	1261.6	40.8	5343	8	Abx76219 Lung canc
20	1261.6	40.8	5343	11	Adn39057 Cancer/an

21	1261.6	40.8	5481	8	Abx76420 Lung canc
22	1261.6	40.8	5481	8	Abx76221 Lung canc
23	1261.6	40.8	5481	11	Adn39061 Cancer/an
24	1261.6	40.8	8146	4	Aah73240 Human cer
25	1251.2	40.5	5363	8	Abx76220 Lung canc
26	1251.2	40.5	5363	8	Abx76419 Lung canc
27	1251.2	40.5	5363	11	Adn39059 Cancer/an
28	1251.2	40.5	5363	11	Adn39727 Cancer/an
29	1250.6	40.5	5481	11	Adn39728 Cancer/an
30	1242.2	40.5	3350	12	Adq18133 Human sof
31	1116.2	36.1	3340	11	Adn39729 Cancer/an
32	1116.2	36.1	4220	11	Adn39731 Cancer/an
33	1114.6	36.1	4347	11	Adn39724 Cancer/an
34	1113	36.0	4480	11	Adn39730 Cancer/an
35	1113	36.0	6924	2	Aaq67283 Human rpt
36	935.2	30.3	6887	12	Adj76009 Marker ge
37	935.2	30.3	6887	12	Ado31209 Mouse pho
38	926.4	30.0	6801	10	Adb79880 Rat phosp
39	926.4	30.0	6801	12	Ado31207 Rat phosp
40	189	6.1	4338	2	Aaq71842 Human rpt
41	189	6.1	6687	12	Adq22975 Human sof
42	60.2	1.9	18817	6	Abi34494 Human met
43	60.2	1.9	18817	6	Abi70161 Chemical
44	60.2	1.9	18817	7	Adg99755 Bisulphit
45	58.4	1.9	6465	6	Abi32985 Human imm

ALIGNMENTS

RESULT 1
AAD48126
ID AAD48126 standard; DNA; 3091 BP.
AC AAD48126;
XX
XX 24-FEB-2003 (first entry)
XX Human protein tyrosine phosphatase zeta SM1 variant DNA.
KW Brain; tumour protein target; Tbt; ischaemic stroke; cancer; epilepsy;
KW schizophrenia; depression; Alzheimer's disease; Parkinson's disease;
KW Huntington's chorea; traumatic head injury; dementia; stupor; headache;
KW coma; vertigo; weakness; myasthenia gravis; cerebrovascular disorder;
KW infection; multiple sclerosis; pregnancy; medical illness; vasotropic;
KW metabolic deficiency; cerebroprotective; antidepressant; antibacterial;
KW cytosolic; nontropic; analgesic; fungicide; virucide; human; enzyme;
KW protein tyrosine phosphatase zeta SM1; gene; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 148..1272
FT /tag= a
FT /product= "Human protein tyrosine phosphatase zeta SM1
FT variant"
FT misc_feature 1262
FT /tag= b
FT /note= "Coding sequence alternative splice site"
FT 3'UTR 1273..3091
FT /tag= c
XX WO200276510-A1.
XX 03-OCT-2002.
XX 22-MAR-2002; 2002WO-US008992.
XX Adn39725 Cancer/an
XX Adn39053 Cancer/an
XX Aa48127 Human pro
XX Adq22702 Human sof
XX Abx76418 Lung canc
XX Abx76219 Lung canc
XX Adn39057 Cancer/an

PI Mueller S, Melcher T, Chin DJ;
XX WPI; 2003-029903/02.
DR P-PSDB; AAE30333.
XX
PT Developing active agents that modulate the activity of a brain tumor
PT protein target gene or gene product for treating e.g. stroke or cancer,
PT comprises contacting an agent with a brain tumor protein.
XX
PS Claim 1; Page 81-84; 135pp; English.
XX
CC The invention relates to a method for developing biologically active
CC agents that modulate activity of a brain tumour protein target (Tbt) gene
CC or gene product. The method is useful for developing biologically active
CC agents that modulate the activity of a brain tumour protein target gene
CC or gene product. Compounds that bind to the brain tumour proteins are
CC useful for treating e.g. ischemic stroke, brain cancer, epilepsy,
CC schizophrenia, depression, Alzheimer's disease, Parkinson's disease,
CC Huntington's chorea, traumatic head injury, dementia, stupor, headache,
CC coma, vertigo, weakness, myasthenia gravis, cerebrovascular disorders,
CC infectious disorders (including fungal, bacterial, viral and parasitic
CC infections), multiple sclerosis, and other complications associated with
CC pregnancy, medical illness, alcohol and substance abuse, toxins and
CC metabolic deficiencies. The brain tumour proteins may also be used to
CC raise antibodies. The present sequence is human protein tyrosine
CC phosphatase beta SM1 variant DNA used to illustrate the method of the
CC invention
XX
SQ Sequence 3091 BP; 1008 A; 535 C; 590 G; 958 T; 0 U; 0 Other;
Query Match 100.0%; Score 3091; DB 8; Length 3091;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CACACATACGACGACGATCTCCTCGATCTATACACTGGAGATTAACAAACAAA 60
Db 1 CACACATACGACGACGATCTCCTCGATCTATACACTGGAGATTAACAAACAAA 60
Qy 61 CAAACAAATTCCTTGGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 120
Db 61 CAAACAAATTCCTTGGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 120
Qy 121 CGGAGAGGGCGGACGCGCTCGAATCGAATCGAATCGAATCGAATCGAATCGAAT 180
Db 121 CGGAGAGGGCGGACGCGCTCGAATCGAATCGAATCGAATCGAATCGAATCGAAT 180
Qy 181 CAGTCCTCTGTGTGGCTCGATGGGCTAATGGATCTACAGACACAGAGAAA 240
Db 181 CAGTCCTCTGTGTGGCTCGATGGGCTAATGGATCTACAGACACAGAGAAA 240
Qy 241 CTTTGAAGAGATGGCTGGCTCTATACAGAGACCTGAATCAAAAATTTGGGAAAG 300
Db 241 CTTTGAAGAGATGGCTGGCTCTATACAGAGACCTGAATCAAAAATTTGGGAAAG 300
Qy 301 AAAATCAACATGTAATAGCCCAACAACTCTCTATCAATATTGATGAAGATCTTACA 360
Db 301 AAAATCAACATGTAATAGCCCAACAACTCTCTATCAATATTGATGAAGATCTTACA 360
Qy 361 CAAATAAATGTGAATCTTAAAGAACTTAAATTTAGGGTTGGGATAAACAATCAATGGAA 420
Db 361 CAAATAAATGTGAATCTTAAAGAACTTAAATTTAGGGTTGGGATAAACAATCAATGGAA 420
Qy 421 AACACATTCATCAACACTGGGAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
Db 421 AACACATTCATCAACACTGGGAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
Qy 481 GTCAGCGAGAGATTCAGAAATGGTGTAAAGCAAGCAAGATAAATTTTCACTGGGGA 540
Db 481 GTCAGCGAGAGATTCAGAAATGGTGTAAAGCAAGCAAGATAAATTTTCACTGGGGA 540
Qy 541 AAATGCAATATGTCACTGTGGATCAGAGCATAGTTTAGAGGACAAAAATTTCCACTT 600
Db 541 AAATGCAATATGTCACTGTGGATCAGAGCATAGTTTAGAGGACAAAAATTTCCACTT 600

Qy 601 GAGATGCAAAATCTACTGCTTTTGTATGCGGACCGATTTTCAAGTTTGGAGAGCAGTCAA 660
Db 601 GAGATGCAAAATCTACTGCTTTTGTATGCGGACCGATTTTCAAGTTTGGAGAGCAGTCAA 660
Qy 661 GGAAGAGGGAGTTTAAGAGCTTTTATCATTTTGTTCAGGTGGGACAGAGAAATTTG 720
Db 661 GGAAGAGGGAGTTTAAGAGCTTTTATCATTTTGTTCAGGTGGGACAGAGAAATTTG 720
Qy 721 GATTTCAAAGCGATTAATTGATGGAGTCCAAAGTGTATTAGTCCTTTTGGGAGCAGGCTGCT 780
Db 721 GATTTCAAAGCGATTAATTGATGGAGTCCAAAGTGTATTAGTCCTTTTGGGAGCAGGCTGCT 780
Qy 781 TTAGATCCATTCATACGTGTTGAACCTTCTGCCAAACTCAACTGACACAGTATTACATTAC 840
Db 781 TTAGATCCATTCATACGTGTTGAACCTTCTGCCAAACTCAACTGACACAGTATTACATTAC 840
Qy 841 AATGGCTCATTTGACATCTCTCTCCCTGCACAGACACAGTGCAGTGGATTGTTTTAAAGAT 900
Db 841 AATGGCTCATTTGACATCTCTCTCCCTGCACAGACACAGTGCAGTGGATTGTTTTAAAGAT 900
Qy 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTTTGTGAAGTTCTTACAATGCAACAA 960
Db 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTTTGTGAAGTTCTTACAATGCAACAA 960
Qy 961 TCTGGTTATGTCATGCTGATGGACTACTTACAAACAAATTTTCGAGAGCACAGTACAAG 1020
Db 961 TCTGGTTATGTCATGCTGATGGACTACTTACAAACAAATTTTCGAGAGCACAGTACAAG 1020
Qy 1021 TTCTCTAGACAGGTTGTTTTCTCTATACACTGGAAAGCAAGATTCATGAAGCAGATTGTT 1080
Db 1021 TTCTCTAGACAGGTTGTTTTCTCTATACACTGGAAAGCAAGATTCATGAAGCAGATTGTT 1080
Qy 1081 AGTTTCAGAACGAGAAATGTTTCAGGCTGACCCAGAGAAATTTATACAGCCCTTTCTGTTACA 1140
Db 1081 AGTTTCAGAACGAGAAATGTTTCAGGCTGACCCAGAGAAATTTATACAGCCCTTTCTGTTACA 1140
Qy 1141 TGGAAAGACCTCGAGTCGTTTATGATACCATGATTCGAGAGTTTCAGTTGTTGATACAG 1200
Db 1141 TGGAAAGACCTCGAGTCGTTTATGATACCATGATTCGAGAGTTTCAGTTGTTGATACAG 1200
Qy 1201 CAGTTGATGGAGAGGACCAACCAAGCATGAATTTTTCAGATGGCTATCAAGACTTG 1260
Db 1201 CAGTTGATGGAGAGGACCAACCAAGCATGAATTTTTCAGATGGCTATCAAGACTTG 1260
Qy 1261 GTAACTATATGATCAGTTGTTTTATAGGGTAACTATATAATTTTCCAAAGGTAAG 1320
Db 1261 GTAACTATATGATCAGTTGTTTTATAGGGTAACTATATAATTTTCCAAAGGTAAG 1320
Qy 1321 AACTTACAAATGGTTGATATATTTTCTCCCTTACCTTTTGTAGACTTTTGTGAGGTTG 1380
Db 1321 AACTTACAAATGGTTGATATATTTTCTCCCTTACCTTTTGTAGACTTTTGTGAGGTTG 1380
Qy 1381 GGTAGGCTGAGTATTTTAAATTTTAAATAAATTTTAAATTTAGAGCTATCTATAATTA 1440
Db 1381 GGTAGGCTGAGTATTTTAAATTTTAAATAAATTTTAAATTTAGAGCTATCTATAATTA 1440
Qy 1441 TGTTTAAAGTTACATTTAAATGATATCATAACTTTTGGCCAACTAATCACTATAGA 1500
Db 1441 TGTTTAAAGTTACATTTAAATGATATCATAACTTTTGGCCAACTAATCACTATAGA 1500
Qy 1501 GTAGATACATATGATCTTATGACTGGAGATCAATTTAGTGTGGCTTTCTTAAAGATTTTCA 1560
Db 1501 GTAGATACATATGATCTTATGACTGGAGATCAATTTAGTGTGGCTTTCTTAAAGATTTTCA 1560
Qy 1561 TTGTAGATATAGTGCAGAACTCTCAGTCCCTGATACATTTTATATTGTGCTTCCATTAC 1620
Db 1561 TTGTAGATATAGTGCAGAACTCTCAGTCCCTGATACATTTTATATTGTGCTTCCATTAC 1620
Qy 1621 GCTATATCAGACAGGAAAGATGAGTAGGGGACATCAAGTCTCTTTGTTGACCAAAA 1680
Db 1621 GCTATATCAGACAGGAAAGATGAGTAGGGGACATCAAGTCTCTTTGTTGACCAAAA 1680

QY 1681 AAATTTTCAGATAACAGCTGGGAAGTCATGATTGGTTCAGAACTTTGGGGATGTAAGAAA 1740
DB |||||
1681 AAATTTTCAGATAACAGCTGGGAAGTCATGATTGGTTCAGAACTTTGGGGATGTAAGAAA 1740
QY 1741 ACATTTCTTACAAAAGATCCACCCCTGCTCCCTCCACAGCGCATGCGAATAAGTAC 1800
DB |||||
1741 ACATTTCTTACAAAAGATCCACCCCTGCTCCCTCCACAGCGCATGCGAATAAGTAC 1800
QY 1801 AGATTCCTCTTGTGGCTCGAGCATGTCACTATTAATCTTGTCTGTGGAGGAGTGTG 1860
DB |||||
1801 AGATTCCTCTTGTGGCTCGAGCATGTCACTATTAATCTTGTCTGTGGAGGAGTGTG 1860
QY 1861 GCATAGATTAGGGTGTAGTTCAGAAAACCTTCATCTGGATGAGTCCAGAAAGTCCCA 1920
DB |||||
1861 GCATAGATTAGGGTGTAGTTCAGAAAACCTTCATCTGGATGAGTCCAGAAAGTCCCA 1920
QY 1921 CTGCAAGTTAAAGGACACTGGACTCTGCACTCAGGCACCTAGAGCTCTGCAAGTCTCTGG 1980
DB |||||
1921 CTGCAAGTTAAAGGACACTGGACTCTGCACTCAGGCACCTAGAGCTCTGCAAGTCTCTGG 1980
QY 1981 AACCTGCATTTAAATAAATAATGCACTATTAAATTAATGTTTTCATATCATGTGGACAAAATGG 2040
DB |||||
1981 AACCTGCATTTAAATAAATAATGCACTATTAAATTAATGTTTTCATATCATGTGGACAAAATGG 2040
QY 2041 ATAAAATTTTGTAACTTTTAAATTCAGTTGCTCGAATATGAGACACAATGACCTGGG 2100
DB |||||
2041 ATAAAATTTTGTAACTTTTAAATTCAGTTGCTCGAATATGAGACACAATGACCTGGG 2100
QY 2101 AAAATCGTGAATAATAGTATTAATAATGTTTATTTTCAATTAATGCTGGAAGATAT 2160
DB |||||
2101 AAAATCGTGAATAATAGTATTAATAATGTTTATTTTCAATTAATGCTGGAAGATAT 2160
QY 2161 TCTATTACTGTTCTTGTGCATATATATGTCAGAAAAGAGATAACTTATGTTGTTCACTTT 2220
DB |||||
2161 TCTATTACTGTTCTTGTGCATATATATGTCAGAAAAGAGATAACTTATGTTGTTCACTTT 2220
QY 2221 TTCAATGCTCTTGTGTTGCAAAATGCCCCCAATTTATTTGTCTAAAATATTAATTTT 2280
DB |||||
2221 TTCAATGCTCTTGTGTTGCAAAATGCCCCCAATTTATTTGTCTAAAATATTAATTTT 2280
QY 2281 GTTGTAGTACTAAATTTATGAATTTGATGAGTCTGCTTAAATAATGAAACTTCTGAAAC 2340
DB |||||
2281 GTTGTAGTACTAAATTTATGAATTTGATGAGTCTGCTTAAATAATGAAACTTCTGAAAC 2340
QY 2341 TAAATCTGATTTTAAAGCAAAAAAAGGCTAGCTTTTCCAGTTCTTCTATAA 2400
DB |||||
2341 TAAATCTGATTTTAAAGCAAAAAAAGGCTAGCTTTTCCAGTTCTTCTATAA 2400
QY 2401 TTCAAAATACCAAGTTTAACTAAGCAACATTCGATTAATTTTCTTGTAGTTTAA 2460
DB |||||
2401 TTCAAAATACCAAGTTTAACTAAGCAACATTCGATTAATTTTCTTGTAGTTTAA 2460
QY 2461 AATAGAGTATTTTCCAGGACAGGAGAAAAGTTTCTAGGAAAGATACCTAGTGTG 2520
DB |||||
2461 AATAGAGTATTTTCCAGGACAGGAGAAAAGTTTCTAGGAAAGATACCTAGTGTG 2520
QY 2521 TTGGTAGTCTATGAGATAACATTTGTATATTAATTAACATCTTTCTTTAGGGTGCTA 2580
DB |||||
2521 TTGGTAGTCTATGAGATAACATTTGTATATTAATTAACATCTTTCTTTAGGGTGCTA 2580
QY 2581 TTCTCAATTAATTTGTACCAATATGAGTTATGTTCTTCCAGATAGTACCATATGCACTA 2640
DB |||||
2581 TTCTCAATTAATTTGTACCAATATGAGTTATGTTCTTCCAGATAGTACCATATGCACTA 2640
QY 2641 ATGGCTTATATGGAATAATACAGGACCACTGATTTGTGCAATGCTTACTGATAATCTCTG 2700
DB |||||
2641 ATGGCTTATATGGAATAATACAGGACCACTGATTTGTGCAATGCTTACTGATAATCTCTG 2700
QY 2701 GTAAGTGCACAGATACATCTATATATTAATCTCAATAAGAGCTAGTTTAAATCTG 2760
DB |||||
2701 GTAAGTGCACAGATACATCTATATATTAATCTCAATAAGAGCTAGTTTAAATCTG 2760
QY 2761 TATGCAATTGATGCTTTCTCTCTATATTTCTTTGGCCAAAAGGCAAGTATTTCTCTTA 2820

DB 2761 TATGCAATTGATGCTTTCTCTCTATATTTCTTTGGCCAAAAGGCAAGTATTTCTCTTA 2820
QY 2821 AGTCTGGATTGCGGGTAATTTTGGGGCATGGGACCAATTTCTCATTTCCAGCAGGTCTG 2880
DB 2821 AGTCTGGATTGCGGGTAATTTTGGGGCATGGGACCAATTTCTCATTTCCAGCAGGTCTG 2880
QY 2881 GTGCCAGACAATAAGTAAACTTATCTCTTAATTAATTTGGAGTTTACCAATTTGTAATAAGAG 2940
DB 2881 GTGCCAGACAATAAGTAAACTTATCTCTTAATTAATTTGGAGTTTACCAATTTGTAATAAGAG 2940
QY 2941 TGACTAAACATATTTATAACATTTGTATAATCACTTAATGAAATTTGCTATGTAATGTT 3000
DB 2941 TGACTAAACATATTTATAACATTTGTATAATCACTTAATGAAATTTGCTATGTAATGTT 3000
QY 3001 GAGACTGTTATTTGGATAATTAAGAGTTGGTTTAAATTTGTTATTTTCTCTTTTTCAG 3060
DB 3001 GAGACTGTTATTTGGATAATTAAGAGTTGGTTTAAATTTGTTATTTTCTCTTTTTCAG 3060
QY 3061 CCCCAGAGCAATTAATGTAAGTAAAGTATACCA 3091
DB 3061 CCCCAGAGCAATTAATGTAAGTAAAGTATACCA 3091

RESULT 2

ABX76417
ID ABX76417 standard; DNA; 5367 Bp.
XX
AC ABX76417;
XX
DT 02-APR-2003 (first entry)
XX
DB Lung cancer-associated polynucleotide #281.
XX
KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
FN WO200286443-A2.
XX
PD 31-OCT-2002.
XX
PF 18-APR-2002; 2002WO-US012476.
XX
PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Aziz N, Murray R;
XX
DR WPI; 2003-093161/08.
XX
PT P-PSDB; ABUS6688.
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
XX for treating lung cancer, by contacting a biological sample from the
XX patient with a polynucleotide that exhibits increased or decreased
XX expression in lung cancer.
PS Claim 22; Page 404-405; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
XX transcript in a cell from a patient, comprising contacting a biological
XX sample from the patient with a polynucleotide that selectively hybridises
XX to a sequence that is at least 80 % identical to a gene that exhibits

increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the invention

SQ Sequence 5367 BP; 1714 A; 1063 C; 1098 G; 1492 T; 0 U; 0 Other;

Query Match 40.9%; Score 1263.2; DB 8; Length 5367;
Best Local Similarity 99.8%; Pred. No. 2.8e-264;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	1	CACACATACGACGACGACGATCTCACTTCGATCTATACACTGGAGGATTAACAAACAA	60
Db	1	CACACATACGACGACGACGATCTCACTTCGATCTATACACTGGAGGATTAACAAACAA	60
Qy	61	CAAAAAAACAATTCCTTGGCTCCGCCCTCTCTCCACTCTGGAAGCAGAGGACCGCA	120
Db	61	CAAAAAAACAATTCCTTGGCTCCGCCCTCTCTCCACTCTGGAAGCAGAGGACCGCA	120
Qy	121	CGGCGAGGGCGGACGCGCTCGGAATGCGAATCTCTTAAGCGTTTCCTCGCTTGCATT	180
Db	121	CGGCGAGGGCGGACGCGCTCGGAATGCGAATCTCTTAAGCGTTTCCTCGCTTGCATT	180
Qy	181	CAGTCTCTCTGTTTGGCTGGATTTGGGCTAATGGATCTACAGACACAGAGAAA	240
Db	181	CAGTCTCTCTGTTTGGCTGGATTTGGGCTAATGGATCTACAGACACAGAGAAA	240
Qy	241	CTTGTGGAAGAGATTGGCTGGTCTCTATACAGAGACCTGAATCAAAAAATTTGGGGAAG	300
Db	241	CTTGTGGAAGAGATTGGCTGGTCTCTATACAGAGACCTGAATCAAAAAATTTGGGGAAG	300
Qy	301	AAATATCAACATGTAATAGCCCAACAACTCTCTATCAATATTTGATGAAGATCTTACA	360
Db	301	AAATATCAACATGTAATAGCCCAACAACTCTCTATCAATATTTGATGAAGATCTTACA	360
Qy	361	CAAGTAATGTCAATCTTAAGAACTTAAATTTTCAAGGTTGGGATAAAACATCTTGGAA	420
Db	361	CAAGTAATGTCAATCTTAAGAACTTAAATTTTCAAGGTTGGGATAAAACATCTTGGAA	420
Qy	421	AACACATTCATCTAACAACCTGGGAAACAGTGGAAATTAATCTCAATGACTACCGT	480
Db	421	AACACATTCATCTAACAACCTGGGAAACAGTGGAAATTAATCTCAATGACTACCGT	480
Qy	481	GTCTGCGGAGGAGTTTCAAGATGGTGTAAAGCAACAGCAAGATACTTTTCACTGGGA	540
Db	481	GTCTGCGGAGGAGTTTCAAGATGGTGTAAAGCAACAGCAAGATACTTTTCACTGGGA	540
Qy	541	AAATGCAATATGTCACTGATGGATCAGAGCATAGTTTGAAGACAAAAATTTCCACTT	600
Db	541	AAATGCAATATGTCACTGATGGATCAGAGCATAGTTTGAAGACAAAAATTTCCACTT	600
Qy	601	GAGATGCAAAATCTACTGTTGATGCGGACCGAATTTTCAAGTTTGGAGAAAGCATCAAA	660
Db	601	GAGATGCAAAATCTACTGTTGATGCGGACCGAATTTTCAAGTTTGGAGAAAGCATCAAA	660
Qy	661	GGAAAAAGGAGTTAAGAGCTTTATCCATTTTGTGAGGTTGGACAGAGAAATTTTG	720
Db	661	GGAAAAAGGAGTTAAGAGCTTTATCCATTTTGTGAGGTTGGACAGAGAAATTTTG	720
Qy	721	GATTTCAAAGCGATTATTGATGGAGTCGAAAGTGTAGTTCGTTTTTGGGAAGCAGCTGCT	780
Db	721	GATTTCAAAGCGATTATTGATGGAGTCGAAAGTGTAGTTCGTTTTTGGGAAGCAGCTGCT	780

Qy	781	TTAGATCCATTATACATCTGTTGAACCTTCTGCCAAACTCAACTGACAGTATTACATTAC	840
Db	781	TTAGATCCATTATACATCTGTTGAACCTTCTGCCAAACTCAACTGACAGTATTACATTAC	840
Qy	841	AATGGCTCATTTGACATCTCTCCCTGCACAGACAGTTCGACTGGATTGTTTTTAAAGAT	900
Db	841	AATGGCTCATTTGACATCTCTCCCTGCACAGACAGTTCGACTGGATTGTTTTTAAAGAT	900
Qy	901	ACAGTTAGCATCTCTGAAAAGCCAGTTTCGCTGTTTTTGTGAAGTTCTTACAATGCAACA	960
Db	901	ACAGTTAGCATCTCTGAAAAGCCAGTTTCGCTGTTTTTGTGAAGTTCTTACAATGCAACA	960
Qy	961	TCTGGTTATGTCTGATGGAGTACTTACAAAACAATTTTCGAGAGACAGTACAG	1020
Db	961	TCTGGTTATGTCTGATGGAGTACTTACAAAACAATTTTCGAGAGACAGTACAG	1020
Qy	1021	TTCTCTAGACAGGTTGTTTTCTCTATACACTGGAAGGAAGATTCATGAAGCAGTTTGT	1080
Db	1021	TTCTCTAGACAGGTTGTTTTCTCTATACACTGGAAGGAAGATTCATGAAGCAGTTTGT	1080
Qy	1081	AGTTCAGAACCCAGAAAATGTTTCAGGCTGACCCAGAGAATTATACAGCCTTCTTGTACA	1140
Db	1081	AGTTCAGAACCCAGAAAATGTTTCAGGCTGACCCAGAGAATTATACAGCCTTCTTGTACA	1140
Qy	1141	TGGAAAGACCTCGAGTCTGTTATGATACCATGATGAGAGTTTCAGTTTGTGACAG	1200
Db	1141	TGGAAAGACCTCGAGTCTGTTATGATGATGATGATGAGAGTTTCAGTTTGTGACAG	1200
Qy	1201	CAGTTGATGGAGAGGACCAAAACCAAGCATGAATTTTTCAGAGATGCTATCAAGACTTG	1260
Db	1201	CAGTTGATGGAGAGGACCAAAACCAAGCATGAATTTTTCAGAGATGCTATCAAGACTTG	1260
Qy	1261	GTAACATAT 1268	
Db	1261	GTTGCTAT 1268	
RESULT 3			
ABX76218			
ID	ABX76218 standard; DNA; 5367 BP.		
XX	AC ABX76218;		
XX	02-APR-2003 (first entry)		
XX	Lung cancer-associated polynucleotide #87.		
XX	Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;		
XX	antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;		
XX	small cell lung cancer; benign lesion; precancerous lesion; bronchitis;		
XX	chronic obstructive pulmonary disease; hypersensitivity pneumonitis;		
XX	interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.		
OS	Unidentified.		
XX	WO200286443-A2.		
XX	31-OCT-2002.		
XX	18-APR-2002; 2002WO-US012476.		
XX	18-APR-2001; 2001US-0284770P.		
XX	10-MAY-2001; 2001US-0290492P.		
XX	09-NOV-2001; 2001US-0339245P.		
XX	13-NOV-2001; 2001US-0350666P.		
XX	29-NOV-2001; 2001US-0334370P.		
XX	12-APR-2002; 2002US-0372246P.		
XX	(EOSB-) EOS BIOTECHNOLOGY INC.		
XX	Aziz N, Murray R;		

DR WPI; 2003-093161/08.
DR P-PSDB; ABU56494.
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
PS Claim 22; Page 258-259; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridizes
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC invention
XX
SQ Sequence 5367 BP; 1714 A; 1063 C; 1098 G; 1492 T; 0 U; 0 Other;
Query Match 40.9%; Score 1263.2; DB 8; Length 5367;
Best Local Similarity 99.8%; Pred. No. 2.8e-264;
Matches 1465; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CACATACGCGACGACGATCTCCTTCGATCTATACACTGAGGATTAACAAACAAA 60
DB 1 CACATACGCGACGACGATCTCCTTCGATCTATACACTGAGGATTAACAAACAAA 60
QY 61 CAAAAAACAATTTCTTGGCTCCGCTCCCTCTCCTCCTCCTCCTCCTCCTCCTCCT 120
DB 61 CAAAAAACAATTTCTTGGCTCCGCTCCCTCTCCTCCTCCTCCTCCTCCTCCTCCT 120
QY 121 CGGCGAGGGCGCGAGCGCTCGGAAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180
DB 121 CGGCGAGGGCGCGAGCGCTCGGAAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180
QY 181 CADCTCCTCTGTGTGTTGGCGCTGGATTTGGGCTAATGGATCTACAGACAACAGAGAAA 240
DB 181 CADCTCCTCTGTGTGTTGGCGCTGGATTTGGGCTAATGGATCTACAGACAACAGAGAAA 240
QY 241 CTGTGTAAGAGATTGGCTGTCCTATACAGAGACCTGAATCAAAAAAATTTGGGGAAG 300
DB 241 CTGTGTAAGAGATTGGCTGTCCTATACAGAGACCTGAATCAAAAAAATTTGGGGAAG 300
QY 301 AAATATCAACATGTAATAGCCCAAAACAACTCTCTATCAATATTTGATGAAGATCTTACA 360
DB 301 AAATATCAACATGTAATAGCCCAAAACAACTCTCTATCAATATTTGATGAAGATCTTACA 360
QY 361 CAAAGTAATGTGAATCTTAAAGAACTTAAATTTCAAGGTTGGGATTAACATCAATTGGAA 420
DB 361 CAAAGTAATGTGAATCTTAAAGAACTTAAATTTCAAGGTTGGGATTAACATCAATTGGAA 420
QY 421 AACACATTCATTATCAACCTGGGAAAACAGTGGAAAATTAATCTCACTAATGACTACCGT 480
DB 421 AACACATTCATTATCAACCTGGGAAAACAGTGGAAAATTAATCTCACTAATGACTACCGT 480
QY 481 GTCCGCGGAGGAGTTTCAGAAATGTTTAAAGCAACGAGATACTTTTCACTGGGGA 540
DB 481 GTCCGCGGAGGAGTTTCAGAAATGTTTAAAGCAACGAGATACTTTTCACTGGGGA 540
QY 541 AAATGCAATATGTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600

DB 541 AAATGCAATATGTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 GAGATGCAAAATCTACTGCTTTGATGCGGACCGATTTTCAAGTTTGGAGAGCAGTCAAA 660
DB 601 GAGATGCAAAATCTACTGCTTTGATGCGGACCGATTTTCAAGTTTGGAGAGCAGTCAAA 660
QY 661 GGAAGAGGGAAGTTTAAAGAGCTTTTATCCATTTTGTGAGGTTGGGACAGAGAAAATTG 720
DB 661 GGAAGAGGGAAGTTTAAAGAGCTTTTATCCATTTTGTGAGGTTGGGACAGAGAAAATTG 720
QY 721 GATTTCAAAGCGATTTATGATGAGTTCGAAAGTGTAGTCTGTTTGGGAGCAGCTGCT 780
DB 721 GATTTCAAAGCGATTTATGATGAGTTCGAAAGTGTAGTCTGTTTGGGAGCAGCTGCT 780
QY 781 TTAGATCCATTCATCTGTTGAACTTCTGCAAACTCACTCACTCACTCACTCACTCACT 840
DB 781 TTAGATCCATTCATCTGTTGAACTTCTGCAAACTCACTCACTCACTCACTCACTCACT 840
QY 841 AATGGCTCATTTGACATCTCTCCCTGACAGACAGTTCAGTTCGATTTGTTTAAAGAT 900
DB 841 AATGGCTCATTTGACATCTCTCCCTGACAGACAGTTCAGTTCGATTTGTTTAAAGAT 900
QY 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGCTTTTGTGAAAGTTCTTCAATGCAACAA 960
DB 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGCTTTTGTGAAAGTTCTTCAATGCAACAA 960
QY 961 TCTGGTTATGTCATGCTGATGAGTACTTACAAAACAAATTTTCGAGAGCAACAGTACAG 1020
DB 961 TCTGGTTATGTCATGCTGATGAGTACTTACAAAACAAATTTTCGAGAGCAACAGTACAG 1020
QY 1021 TTCTCTAGACAGGTTGTTTCTCTCATACACTGGAAGGAGAGATTTCAAGAGCAGTTTGT 1080
DB 1021 TTCTCTAGACAGGTTGTTTCTCTCATACACTGGAAGGAGAGATTTCAAGAGCAGTTTGT 1080
QY 1081 AGTTCAGAACCCAGAAAATGTTTCAGGTCGACCCAGAGAAATATATCCAGCCTTCTCTGTACA 1140
DB 1081 AGTTCAGAACCCAGAAAATGTTTCAGGTCGACCCAGAGAAATATATCCAGCCTTCTCTGTACA 1140
QY 1141 TGGGAAGACCTCGAGTGGTTTATGATACCATGATGAGAGTTGAGAGTTTGTACAG 1200
DB 1141 TGGGAAGACCTCGAGTGGTTTATGATACCATGATGAGAGTTGAGAGTTTGTACAG 1200
QY 1201 CAGTTGGAATGAGAGGAGGACCAACCAAGCATCAATTTTTTGACAGATGGCTATCAAGACTTG 1260
DB 1201 CAGTTGGAATGAGAGGAGGACCAACCAAGCATCAATTTTTTGACAGATGGCTATCAAGACTTG 1260
QY 1261 GTAACTAT 1268
DB 1261 GGTGCTAT 1268
RESULT 4
ADN39055
ID ADN39055 standard; cDNA; 5367 BP.
XX
AC ADN39055;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:373.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.

XX PD 22-MAY-2003.
XX PF 13-NOV-2002; 2002WO-US036810.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 21-NOV-2001; 2001US-0332464P.
XX PR 29-NOV-2001; 2001US-0334393P.
XX PR 03-DEC-2001; 2001US-0335394P.
XX PR 14-DEC-2001; 2001US-0340376P.
XX PR 08-JAN-2002; 2002US-0347211P.
XX PR 10-JAN-2002; 2002US-0347349P.
XX PR 08-FEB-2002; 2002US-0355250P.
XX PR 13-FEB-2002; 2002US-0356714P.
XX PR 20-FEB-2002; 2002US-0359077P.
XX PR 29-MAR-2002; 2002US-0368809P.
XX PR 04-APR-2002; 2002US-0370110P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PR 05-JUN-2002; 2002US-0386614P.
XX PR 16-JUL-2002; 2002US-0396839P.
XX PR 22-JUL-2002; 2002US-0397775P.
XX PR 22-JUL-2002; 2002US-0397845P.
XX PR 09-SEP-2002; 2002US-0409450P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PF Afar D, Aziiz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilison KE, Zlotnik A;
XX P-PSDB; ADN39056.
XX WPI; 2003-468649/44.
XX P-PSDB; ADN39056.
XX Determining the presence or absence of a pathological cell in a patient,
XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.
XX Claim 8; SEQ ID NO 373; 1385pp; English.
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
XX whose expression is upregulated or downregulated in specific cancers or
XX other diseases such as angiotensin or fibrotic disorders, and to methods
XX of determining the presence or absence of a pathological cell in a
XX patient by detecting a nucleic acid at least 80% identical to those of
XX the invention or by detecting a polypeptide of the invention. The
XX invention also relates to expression vectors and host cells comprising a
XX nucleic acid of the invention; antibodies which specifically bind a
XX polypeptide of the invention; use of such antibodies for drug targeting;
XX and methods of screening for modulators of activity or expression of the
XX polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX antibodies and methods are useful for diagnosing, prognosing and treating
XX cancer and other conditions such as psoriasis, ischaemia, heart disease,
XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX neovascularization syndromes, scarring and uterine fibroids. They may
XX also be useful in wound healing and in contraception. The present
XX sequence represents a nucleic acid sequence of the invention.
XX Sequence 5367 BP; 1714 A; 1063 C; 1098 G; 1492 T; 0 U; 0 Other;
Query Match 40.9%; Score 1263.2; DB 11; Length 5367;
Best Local Similarity 99.8%; Pred. No. 2.8e-264;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CACACATCGGCGCAGCATCTCACTTCGATCTATACACTGGAGGATTAAACACAA 60
DB 1 CACACATCGGCGCAGCATCTCACTTCGATCTATACACTGGAGGATTAAACACAA 60
QY 61 CAAAAAACAATTTCTTGGCTCCCTCCCTCTCCACTCTGAGAAGCAGAGGCCGCA 120
DB 61 CAAAAAACAATTTCTTGGCTCCCTCCCTCTCCACTCTGAGAAGCAGAGGCCGCA 120
QY 121 CGGCGAGGGGGCGCAGACCGCTTGAAATGCGAATCCCTAAAGCGTTTCTCGCTTGCATT 180
DB 121 CGGCGAGGGGGCGCAGACCGCTTGAAATGCGAATCCCTAAAGCGTTTCTCGCTTGCATT 180

QY 181 CAGCTCCTCTGTGTTTGC CGCCTGGATTGGGCTAATCGATACTACAGACAACAGAGAAA 240
DB 181 CAGCTCCTCTGTGTTTGC CGCCTGGATTGGGCTAATCGATACTACAGACAACAGAGAAA 240
QY 241 CTTGTTGAAGAGATTGGCTGCTCCTATACAGAGGACACTGAATCAAAAAAATTTGGGGAAAG 300
DB 241 CTTGTTGAAGAGATTGGCTGCTCCTATACAGAGGACACTGAATCAAAAAAATTTGGGGAAAG 300
QY 301 AAATATCCAACTAGTAATAGCCAAAAACAATCTCTCTATCAATATTGATGAAGATCTTACA 360
DB 301 AAATATCCAACTAGTAATAGCCAAAAACAATCTCTCTATCAATATTGATGAAGATCTTACA 360
QY 361 CAAGTAAATGTGAATCTTAAAGAACTTTAAATTTTCAGGGTTGGGATAAAACATCAITGGAA 420
DB 361 CAAGTAAATGTGAATCTTAAAGAACTTTAAATTTTCAGGGTTGGGATAAAACATCAITGGAA 420
QY 421 AACACATTCATTCAATAACACTGGGAAAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
DB 421 AACACATTCATTCAATAACACTGGGAAAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
QY 481 GTCAGCGGAGGAGTTTCAGAAATGGTGTTTAAAGCAAGCAAGATAAATTTCTACTGGGGA 540
DB 481 GTCAGCGGAGGAGTTTCAGAAATGGTGTTTAAAGCAAGCAAGATAAATTTCTACTGGGGA 540
QY 541 AAATGCAATATGTCATCTGATGGATCAGAGCATAGTTTAGAAGGACAAAAAATTTCCACTT 600
DB 541 AAATGCAATATGTCATCTGATGGATCAGAGCATAGTTTAGAAGGACAAAAAATTTCCACTT 600
QY 601 GAGATGCAAAATCTACTGCTTTGATGCGGACCGAATTTTCAAGTTTGTAGGAGCAGTCAAA 660
DB 601 GAGATGCAAAATCTACTGCTTTGATGCGGACCGAATTTTCAAGTTTGTAGGAGCAGTCAAA 660
QY 661 GGAAGGAGGAGTTTAAGAGCTTTTATCCATTTTGTGAGGTTGGGACAGAGAAATTTG 720
DB 661 GGAAGGAGGAGTTTAAGAGCTTTTATCCATTTTGTGAGGTTGGGACAGAGAAATTTG 720
QY 721 GATTTCAAGCGGATTATTGATGGAGTCGAAAGTCGTTAGTCTGTTTGGGAGCAGGCTGT 780
DB 721 GATTTCAAGCGGATTATTGATGGAGTCGAAAGTCGTTAGTCTGTTTGGGAGCAGGCTGT 780
QY 781 TTAGATCCATTCATPACTGTTGAACCTTCTGCCAACTCAACTGCAAGATTTACATTTAC 840
DB 781 TTAGATCCATTCATPACTGTTGAACCTTCTGCCAACTCAACTGCAAGATTTACATTTAC 840
QY 841 AATGGCTCATTTGACATCTCTCCCTGCACAGACAGTTCGATTCGATTTGTTTAAAGAT 900
DB 841 AATGGCTCATTTGACATCTCTCCCTGCACAGACAGTTCGATTCGATTTGTTTAAAGAT 900
QY 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTGTGNAAGTTCTCAATGCAACNA 960
DB 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTGTGNAAGTTCTCAATGCAACNA 960
QY 961 TCTGTTTATGTCATGCTGATGGACTACTTACAAACCAATTTTCGAGAGCAACAGTACAAG 1020
DB 961 TCTGTTTATGTCATGCTGATGGACTACTTACAAACCAATTTTCGAGAGCAACAGTACAAG 1020
QY 1021 TTCTCTAGACAGGTTGTTTCTCTATACACTGGAAGGAGGATTCATGAGCAGTTTGT 1080
DB 1021 TTCTCTAGACAGGTTGTTTCTCTATACACTGGAAGGAGGATTCATGAGCAGTTTGT 1080
QY 1081 AGTTTCAGAACCCAGAAAAATGTTTCAGGCTGACCCAGAGAAATATATACAGCCTTCTTGTTACA 1140
DB 1081 AGTTTCAGAACCCAGAAAAATGTTTCAGGCTGACCCAGAGAAATATATACAGCCTTCTTGTTACA 1140
QY 1141 TGGGAAAGACCTCGAGTCGTTTATGATACCAATGATGAGAGTTTGAGATTTGATCAAG 1200
DB 1141 TGGGAAAGACCTCGAGTCGTTTATGATACCAATGATGAGAGTTTGAGATTTGATCAAG 1200
QY 1201 CAGTTGATGAGAGGAGGACCAACCAAGCATGAAATTTTTCAGAGATGGCTATCAAGACTTG 1260
DB 1201 CAGTTGATGAGAGGAGGACCAACCAAGCATGAAATTTTTCAGAGATGGCTATCAAGACTTG 1260

Db 841 AATGGCTCATTTGACATCTCCTCCCTGCACACACAGTTGACTGGATTGTTTTAAAGAT 900
Qy 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTTTGTGAAGTCTTCAATGCAACAA 960
Db 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTTTGTGAAGTCTTCAATGCAACAA 960
Qy 961 TCTGTTATGTCATGCTGATGGACTACTTACAAACAAATTTTCAGAGAGCAACAGTACAAG 1020
Db 961 TCTGTTATGTCATGCTGATGGACTACTTACAAACAAATTTTCAGAGAGCAACAGTACAAG 1020
Qy 1021 TTCTCTAGACAGGTGTTTTCTCTCATACACTGGAAGGAAGATTCATGAAGCAGTTTGT 1080
Db 1021 TTCTCTAGACAGGTGTTTTCTCTCATACACTGGAAGGAAGATTCATGAAGCAGTTTGT 1080
Qy 1081 AGTTCAGAACAGAAATGTTTCAGCTGACCCAGCAGAAATATACAGCCTCTTGTGTACA 1140
Db 1081 AGTTCAGAACAGAAATGTTTCAGCTGACCCAGCAGAAATATACAGCCTCTTGTGTACA 1140
Qy 1141 TGGGAAGACCTCGAGTCGTTTTATGATACCATGATTGAGAAAGTTTGCAGTTTTTGTACCAG 1200
Db 1141 TGGGAAGACCTCGAGTCGTTTTATGATACCATGATTGAGAAAGTTTGCAGTTTTTGTACCAG 1200
Qy 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGAATTTTTCAGAGATGGCTATCAAGACTTG 1260
Db 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGAATTTTTCAGAGATGGCTATCAAGACTTG 1260
Qy 1261 GTAACATAT 1268
Db 1261 GGTGCTAT 1268

RESULT 6

ABL66502
ID ABL66502 standard; DNA; 7941 BP.
XX
AC ABL66502;
XX
DT 15-MAY-2002 (first entry)
DE Lung cancer related gene sequence SEQ ID NO:4839.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cycostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.

PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
PA (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.

Claim 1; SEQ ID NO 4839; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's tumour

Sequence 7941 BP; 2406 A; 1651 C; 1598 G; 2286 T; 0 U; 0 Other;

Query Match 40.9%; Score 1263.2; DB 6; Length 7941;
Best Local Similarity 99.8%; Pred. No. 3.1e-264;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CACACATACGCACGACGATCTCACTTCGATCTATACACTGGAGGATTTAAACAAACAAA 60
Db 1 CACACATACGCACGACGATCTCACTTCGATCTATACACTGGAGGATTTAAACAAACAAA 60
Qy 61 CAAAAAAAACATTTTCCTTCGCTCCCTCCCTCTCCACTCTGAGAAGCAGAGCGCGCA 120
Db 61 CAAAAAAAACATTTTCCTTCGCTCCCTCCCTCTCCACTCTGAGAAGCAGAGCGCGCA 120
Qy 121 CGCGAGGGCGCGAGACCGCTTCGAAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180
Db 121 CGCGAGGGCGCGAGACCGCTTCGAAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180

181 QY CAGCTCTCTGTGTTTGGCCGCTGGATGGCTTAATGGCTACTACAGACAACAGAGAAA 240
181 Db |||||
241 QY CTGTTGAAGAGATTGGCTGGCTCTTATACAGAGAGCACTGAATCAAAAAATTTGGGGAAG 300
241 Db |||||
301 QY AAATATCCAACTGTATATAGCCMAAACAACTCTCTATCAATATTTGATGAAGATCTTACA 360
301 Db |||||
361 QY CAAGTAAATGTGAATCTTAAGAACTTAATTTTCAGGGTTGGGATAAAACATCATTTGGA 420
361 Db |||||
421 QY AACACATTTCAATCACTGGAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
421 Db |||||
481 QY GTGAGCGGAGGAGTTTCAGAAATGGTGTAAAGCAAGCAAGATACTTTTCACTGGGA 540
481 Db |||||
541 QY AAATGCAATATGTCACTGATGATCAGAGCATAGTTTAAAGAGCAAAAAATTTCCACTT 600
541 Db |||||
601 QY GAGTGCAAATCTACTGCTTTGATGCGGACCGATTTTCAAGTTTTCAGGAAGCACTCAA 660
601 Db |||||
661 QY GCAAAAGGGAAGTAAAGACTTTATCCATTTTGTGAGTTGGGACAGAAATTTG 720
661 Db |||||
721 QY GATTTCAAAGCGATTAATGATGAGTTCGAAAGTGTGTAGTGGGAGCAGGCTGCT 780
721 Db |||||
781 QY TTAGATCCATTCATCTGTGTGAACCTTCTGCCAACTCACTGACAAATTTACATTTAC 840
781 Db |||||
841 QY AATGCTCATTCACATCTCTCCCTGCACAGACAGATTTGACTGATTTTAAAGAT 900
841 Db |||||
901 QY ACGTTAGCATCTGAAAGCCAGTTGGCTGTTTTTGTGAAGTTCTTACAATGCAACA 960
901 Db |||||
961 QY TCTGTTATGTCATGCTGATGACTACTTACAAACAAATTTTCGAGAGCAACAGTACAG 1020
961 Db |||||
1021 QY TTCTCTAGACAGTGTTCCTCTATACCTGAAAGGAGAGATTCATGAAGCACTTTGT 1080
1021 Db |||||
1081 QY AGTTTCAGAACCAAGAAATGTTTCAGGCTGACCCAGAGAAATTTATACCAGCTTCTTTTACA 1140
1081 Db |||||
1141 QY TGGGAAAGACCTCGAGTCTGTTTATGATACCATGATGAGAGTTTTCAGTTTGTACAG 1200
1141 Db |||||
1201 QY CAGTTGGATGGAGAGGACCAACCAAGCATGAATTTTGGACAGATGGCTATCAAGACTTG 1260
1201 Db |||||
1261 QY GAACTAT 1268

1261 GGTGCTAT 1268
RESULT 7
ABX76421
ID ABX76421 standard; DNA; 7941 BP.
XX AC ABX76421;
XX DT 02-APR-2003 (first entry)
XX DE Lung cancer-associated polynucleotide #285.
XX KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
XX KW antiinflammatory; antisthmatic; non-small cell lung cancer; atelectasis;
XX KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
XX KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
XX KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX OS Unidentified.
XX PN WO200286443-A2.
XX PD 31-OCT-2002.
XX PF 18-APR-2002; 2002WO-US012476.
XX PR 18-APR-2001; 2001US-0284770P.
XX PR 10-MAY-2001; 2001US-0290492P.
XX PR 09-NOV-2001; 2001US-0339245P.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 29-NOV-2001; 2001US-0343370P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Aziz N, Murray R;
XX DR WPI; 2003-093161/08.
XX DR P-PSDB; ABUS6692.
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
PS Claim 22; Page 410-412; 453pp; English.
XX The invention relates to a method for detecting a lung cancer-associated
XX transcript in a cell from a patient, comprising contacting a biological
XX sample from the patient with a polynucleotide that selectively hybridizes
XX to a sequence that is at least 80 % identical to a gene that exhibits
XX increased or decreased expression in lung cancer samples. Lung cancer-
XX associated polynucleotides and polypeptides are used for identifying a
XX compound that modulates a lung cancer-associated polypeptide, for
XX inhibiting proliferation of a lung cancer-associated cell to treat lung
XX cancer in a patient and for treating a mammal having lung cancer by
XX administering a modulatory compound identified. The methods are useful
XX for treating lung cancer, such as small cell lung cancer, non-small cell
XX lung cancer or other benign or precancerous lesions, e.g. atelectasis,
XX emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
XX hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
XX bronchiectasis. The genes, polynucleotides and polypeptides are useful
XX for diagnostic purposes and as targets for screening for therapeutic
XX compounds that modulate lung cancer, such as antibodies. Sequences
XX ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
XX invention
SQ Sequence 7941 BP; 2406 A; 1651 C; 1598 G; 2286 T; 0 U; 0 Other;
Query Match 40.9%; Score 1263.2; DB 8; Length 7941;
Best Local Similarity 99.8%; Pred. No. 3.1e-264;

agents that modulate activity of a brain tumour protein target (Tbt) gene or gene product. The method is useful for developing biologically active agents that modulate the activity of a brain tumour protein target gene or gene product. Compounds that bind to the brain tumour proteins are useful for treating e.g. ischaemic stroke, brain cancer, epilepsy, schizophrenia, depression, Alzheimer's disease, Parkinson's disease, Huntington's chorea, traumatic head injury, dementia, stupor, headache, coma, vertigo, weakness, myasthenia gravis, cerebrovascular disorders, infectious disorders (including fungal, bacterial, viral and parasitic infections), multiple sclerosis, and other complications associated with pregnancy, medical illnesses, alcohol and substance abuse, toxins and metabolic deficiencies. The brain tumour proteins may also be used to raise antibodies. The present sequence is human protein tyrosine phosphatase receptor-type 2 polypeptide 1 (PTPRZ1) encoding DNA used to illustrate the method of the invention

Sequence 7941 BP; 2407 A; 1651 C; 1597 G; 2286 T; 0 U; 0 Other;

Query Match 40.9%; Score 1263.2; DB 8; Length 7941;
Best Local Similarity 99.8%; Pred. No. 3.1e-264;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACACATACGACGACGATCTCCTTCGATCTATACACTGAGGATTAACCAACAAA 60
Db 1 CACACATACGACGACGATCTCCTTCGATCTATACACTGAGGATTAACCAACAAA 60
QY 61 CAAAAAAACATTTCTTCGCTCCCTCCCTCTCCACTCTGAGAGCAGAGCCGCA 120
Db 61 CAAAAAAACATTTCTTCGCTCCCTCCCTCTCCACTCTGAGAGCAGAGCCGCA 120
QY 121 CGGCGAGGGCCGAGACCGCTCGGAATGCGAATCTTAAAGCGTTCTCGCTTGCATT 180
Db 121 CGGCGAGGGCCGAGACCGCTCGGAATGCGAATCTTAAAGCGTTCTCGCTTGCATT 180
QY 181 CAdCTCTCTGTTGTTGGCGCTGATTTGGGCTTAATGATCTACAGACCAACAGAGAAA 240
Db 181 CAGCTCTCTGTTGTTGGCGCTGATTTGGGCTTAATGATCTACAGACCAACAGAGAAA 240
QY 241 CTGTGTGAAGAGATTTGGCTGCTTATACAGGACGACCTGAATCAAAAAATTTGGGGAAG 300
Db 241 CTGTGTGAAGAGATTTGGCTGCTTATACAGGACGACCTGAATCAAAAAATTTGGGGAAG 300
QY 301 AAATATCCAAATGATAGCCCAAAACAAATCTCTATCAATATGATGAAGATCTTACA 360
Db 301 AAATATCCAAATGATAGCCCAAAACAAATCTCTATCAATATGATGAAGATCTTACA 360
QY 361 CAAGTAATGTAATCTTAAGAACTTAATTTCAAGGCTTGGGATTAACATCATTTGAA 420
Db 361 CAAGTAATGTAATCTTAAGAACTTAATTTCAAGGCTTGGGATTAACATCATTTGAA 420
QY 421 AACACATTCATTAACAACATGGGAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
Db 421 AACACATTCATTAACAACATGGGAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
QY 481 GTACGCGGAGGATTTCAAGATGTTTAAAGCAAGCAAGATACTTTTCACTGGGGA 540
Db 481 GTACGCGGAGGATTTCAAGATGTTTAAAGCAAGCAAGATACTTTTCACTGGGGA 540
QY 541 AAATGCAATATGTCATCTGATGATCAGGATAGTTTGAAGGACAAATTTCCACTT 600
Db 541 AAATGCAATATGTCATCTGATGATCAGGATAGTTTGAAGGACAAATTTTCCACTT 600
QY 601 GAGATGCAAACTACTGCTTGTATCGGACCGAATTTTCAAGTTTGGAGGACGAGTCAAA 660
Db 601 GAGATGCAAACTACTGCTTGTATCGGACCGAATTTTCAAGTTTGGAGGACGAGTCAAA 660
QY 661 GGAAGGGAAGTTTAAGAGCTTTATCCATTTTGTGAGGTTGGACAGAGAAATTTG 720
Db 661 GGAAGGGAAGTTTAAGAGCTTTATCCATTTTGTGAGGTTGGACAGAGAAATTTG 720
QY 721 GATTTCAAGCGATTTATGAGGATTCGAAAGTGTAGTCTGTTTGGAGCAGCTGCT 780
Db 721 GATTTCAAGCGATTTATGAGGATTCGAAAGTGTAGTCTGTTTGGAGCAGCTGCT 780

QY 781 TTAGATCCATTCACTACTGTTGAACCTTCTGCCAAACTCAACTGACAAGTATTACATTAC 840
Db 781 TTAGATCCATTCACTACTGTTGAACCTTCTGCCAAACTCAACTGACAAGTATTACATTAC 840
QY 841 AATGGCTCATTTGACATCTCTCCCTGCAACAGATTTGATTTGTTTAAAGAT 900
Db 841 AATGGCTCATTTGACATCTCTCCCTGCAACAGATTTGATTTGTTTAAAGAT 900
QY 901 ACAGTTAGCATCTCTGAAGCCAGTTGGCTGCTTTTGTGAAGTCTTACATGCAACAA 960
Db 901 ACAGTTAGCATCTCTGAAGCCAGTTGGCTGCTTTTGTGAAGTCTTACATGCAACAA 960
QY 961 TCTGTTATGTCATGCTGATGAGTCTTACAAAACAATTTTCGAGAGCAACAGTACAA 1020
Db 961 TCTGTTATGTCATGCTGATGAGTCTTACAAAACAATTTTCGAGAGCAACAGTACAA 1020
QY 1021 TTCTCTAGACAGGTTGTTTCTCTATACACTGCGAAAGGAGATTTTCATGAAGCAGTTGT 1080
Db 1021 TTCTCTAGACAGGTTGTTTCTCTATACACTGCGAAAGGAGATTTTCATGAAGCAGTTGT 1080
QY 1081 AGTTTCAGAACCCAGAAAATGTTTCAGGCTGACCCAGAGATTTATACAGGCTTCTTGT 1140
Db 1081 AGTTTCAGAACCCAGAAAATGTTTCAGGCTGACCCAGAGATTTATACAGGCTTCTTGT 1140
QY 1141 TGGGAAGACCTCGAGTCTGTTTATGATACCATGATTTGAGAAGTTTGACAGTTTGTAC 1200
Db 1141 TGGGAAGACCTCGAGTCTGTTTATGATACCATGATTTGAGAAGTTTGACAGTTTGTAC 1200
QY 1201 CAGTTGATGAGAGAGGACCAAAACCAAGCATGAATTTTTCAGAGATGGCTATCAAGATTG 1260
Db 1201 CAGTTGATGAGAGAGGACCAAAACCAAGCATGAATTTTTCAGAGATGGCTATCAAGATTG 1260
QY 1261 GTAACTAT 1268
Db 1261 GGTGCTAT 1268

RESULT 9

ACC72641
ID ACC72641 standard; cDNA; 7941 BP.
XX
AC ACC72641;
XX
DT 09-JUL-2003 (first entry)
XX
DE Human protein tyrosine phosphatase receptor-type 2 polypeptide 1 cDNA.
XX
KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
XX
KW heart disease; atherosclerosis; endometriosis; gene; ss.
XX
OS Homo sapiens.
XX
FN WO2003025138-A2.
XX
PD 27-MAR-2003.
XX
PP 17-SEP-2002; 2002WO-US029560.
XX
PR 17-SEP-2001; 2001US-0323469P.
PR 20-SEP-2001; 2001US-0323887P.
PR 13-NOV-2001; 2001US-0350666P.
PR 08-FEB-2002; 2002US-0355145P.
PR 08-FEB-2002; 2002US-0355257P.
PR 12-APR-2002; 2002US-0372246P.
XX
FA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KB;
PI Zlotnik A;
XX
DR WPI; 2003-354600/33.
P-PSDB; ABR58521.

XX New genes that are up-regulated or down-regulated in cancers, useful as
PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
PT therapeutic targets for screening drugs for treating these diseases.
XX
PS Claim 8; Page 138-139; 767pp; English.
XX
CC The present invention describes an isolated nucleic acid molecule, which
CC comprises the sequence of any of the genes that are up-regulated or down-
CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
CC related gene nucleotide sequences which encode the proteins given in
CC ABR58521 to ABR58709. Also described: (1) determining the presence or
CC absence of a pathological cell in a patient; (2) an expression vector
CC comprising a nucleic acid molecule described above; (3) a host cell
CC comprising the vector; (4) an isolated polypeptide, which is encoded by
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
CC of (4); (6) specifically targeting a compound to a pathological cell in a
CC patient by administering to the patient the antibody above; and (7) a
CC drug screening assay. The nucleic acid is useful as diagnostic markers or
CC therapeutic targets. In particular, the nucleic acid is useful for
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
CC atherosclerosis and endometriosis. The nucleic acid is also useful in
CC drug screening, particularly for identifying agents for treating these
CC pathologies
XX
SQ Sequence 7941 BP; 2406 A; 1651 C; 1598 G; 2286 T; 0 U; 0 Other;
Query Match 40.9%; Score 1263.2; DB 10; Length 7941;
Best Local Similarity 99.8%; Pred. No. 3.1e-264;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CACACATACGACGACGATCTCACTTCGATCTATACACTGGAGGATTAACAAACAAA 60
Db 1 CACACATACGACGACGATCTCACTTCGATCTATACACTGGAGGATTAACAAACAAA 60
Qy 61 CAAACAAAACATTTCTCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 120
Db 61 CAAACAAAACATTTCTCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 120
Qy 121 CGGCGAGGGCGGACGCGCTCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGA 180
Db 121 CGGCGAGGGCGGACGCGCTCGAATCGAATCGAATCGAATCGAATCGAATCGAATCGA 180
Qy 181 CAGCTCCTCTGTTGTCGCTGATTTGGCTAATGGATTAATGGATTAATGGATTAATGG 240
Db 181 CAGCTCCTCTGTTGTCGCTGATTTGGCTAATGGATTAATGGATTAATGGATTAATGG 240
Qy 241 CTTGTTGAAGAGATTTGGCTGCTCTATACAGGACACTGAATCAAAAAAATTTGGGAA 300
Db 241 CTTGTTGAAGAGATTTGGCTGCTCTATACAGGACACTGAATCAAAAAAATTTGGGAA 300
Qy 301 AAATATCAACATGTAATAGCCCAACAACTCTCTATCAATATTTGATGAAGATCTTACA 360
Db 301 AAATATCAACATGTAATAGCCCAACAACTCTCTATCAATATTTGATGAAGATCTTACA 360
Qy 361 CAAAGTAAATGTGAATCTTTAAGAACTTAAATTTAGGGTTGGGATAAAACATCAATTG 420
Db 361 CAAAGTAAATGTGAATCTTTAAGAACTTAAATTTAGGGTTGGGATAAAACATCAATTG 420
Qy 421 AACACATTCATTAACACATCGGGAACACAGTGGAAATTAATCTCAATGACTACCGT 480
Db 421 AACACATTCATTAACACATCGGGAACACAGTGGAAATTAATCTCAATGACTACCGT 480
Qy 481 GTCAGCGGAGGAGTTTCAGAAATGTTGTTTAAAGCAAGCAAGATACTTTTCACTGG 540
Db 481 GTCAGCGGAGGAGTTTCAGAAATGTTGTTTAAAGCAAGCAAGATACTTTTCACTGG 540
Qy 541 AAATGCAATATGTCAATCTGATGGATCAGAGCATAGTTTGAAGGACAAATAATTTCC 600
Db 541 AAATGCAATATGTCAATCTGATGGATCAGAGCATAGTTTGAAGGACAAATAATTTCC 600

Qy 601 GAGATGCAANTCTACTCTTCTTGGATGCGGACCGGATTTTCAAGTTTGTGAGGAGCAGTCAA 660
Db 601 GAGATGCAANTCTACTCTTCTTGGATGCGGACCGGATTTTCAAGTTTGTGAGGAGCAGTCAA 660
Qy 661 GGAAGAGGGAAGTAAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAAATTTG 720
Db 661 GGAAGAGGGAAGTAAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAAATTTG 720
Qy 721 GATTTCAAACGCGATTAATGATGAGGTGCGAAGTGTGTAGTCTGTTTGGGAGCAGGCTGCT 780
Db 721 GATTTCAAACGCGATTAATGATGAGGTGCGAAGTGTGTAGTCTGTTTGGGAGCAGGCTGCT 780
Qy 781 TTAGATCCATTCATCTACTGTTGAACTTCTGCCAACTCAACTGACAGATTTACATTTAC 840
Db 781 TTAGATCCATTCATCTACTGTTGAACTTCTGCCAACTCAACTGACAGATTTACATTTAC 840
Qy 841 AATGGCTCATTTGACATCTCTCTCCCTGCACAGACACAGTTCGACTGGATTTGTTTAAAGAT 900
Db 841 AATGGCTCATTTGACATCTCTCTCCCTGCACAGACACAGTTCGACTGGATTTGTTTAAAGAT 900
Qy 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTTTGTGAAAGTTCTTACAATGCAACAA 960
Db 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTTTGTGAAAGTTCTTACAATGCAACAA 960
Qy 961 TCTGGTTATGCTCATGCTGATGGACTACTTACAAAACAAATTTTCGAGAGCAGCAGTACAAG 1020
Db 961 TCTGGTTATGCTCATGCTGATGGACTACTTACAAAACAAATTTTCGAGAGCAGCAGTACAAG 1020
Qy 1021 TTCTCTAGACAGGTTGTTTTCTCTCATACACTGGAAGGAGAGATTCATGAAGCAGATTTGT 1080
Db 1021 TTCTCTAGACAGGTTGTTTTCTCTCATACACTGGAAGGAGAGATTCATGAAGCAGATTTGT 1080
Qy 1081 AGTTTCAGAACCCAGAAAATGTTTCAGGCTGACCCAGAGAAATTATACCGCTTCTTGTTACA 1140
Db 1081 AGTTTCAGAACCCAGAAAATGTTTCAGGCTGACCCAGAGAAATTATACCGCTTCTTGTTACA 1140
Qy 1141 TGGGAAAGACTCGAGTCTGTTTATGATACCATGATTCAGAGAGTTTGCAGTTTGTACCG 1200
Db 1141 TGGGAAAGACTCGAGTCTGTTTATGATACCATGATTCAGAGAGTTTGCAGTTTGTACCG 1200
Qy 1201 CAGTTGATGGAGAGGAGGACCAACCAACATGAAATTTTTCAGAGATGGCTATCAAGACTTG 1260
Db 1201 CAGTTGATGGAGAGGAGGACCAACCAACATGAAATTTTTCAGAGATGGCTATCAAGACTTG 1260
Qy 1261 GTAACTAT 1268
Db 1261 GGTGCTAT 1268
RESULT 10
ADN39063
ID ADN39063 standard; cDNA; 7941 BP.
XX ADN39063;
XX
XX 17-JUN-2004 (first entry)
XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:381.
XX Human; differential expression; cancer; angiogenic disorder;
XX fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
XX inflammatory disease; autoimmune disease;
XX retinal neovascularisation syndrome; scarring; uterine fibroid;
XX detection; diagnosis; prognosis; drug screening; drug targeting;
XX wound healing; contraception; cytostatic; cardiant; immunomodulatory;
XX vulnery; gene therapy; vaccine; gene; ss.
XX Homo sapiens.
XX
XX WO2003042661-A2.
XX
XX 22-MAY-2003.

XX 13-NOV-2002; 2002WO-US036810.
XX 13-NOV-2001; 2001US-0350666P.
XX 21-NOV-2001; 2001US-0332464P.
XX 29-NOV-2001; 2001US-0334393P.
XX 03-DEC-2001; 2001US-0335394P.
XX 14-DEC-2001; 2001US-0340376P.
XX 08-JAN-2002; 2002US-0347211P.
XX 10-JAN-2002; 2002US-0347349P.
XX 08-FEB-2002; 2002US-0352505P.
XX 13-FEB-2002; 2002US-0356714P.
XX 20-FEB-2002; 2002US-0359077P.
XX 29-MAR-2002; 2002US-0368099P.
XX 04-APR-2002; 2002US-0370110P.
XX 12-APR-2002; 2002US-0372246P.
XX 05-JUN-2002; 2002US-0386614P.
XX 16-JUL-2002; 2002US-0396839P.
XX 22-JUL-2002; 2002US-0397755P.
XX 22-JUL-2002; 2002US-0397845P.
XX 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
XX Mack DH, Murray R, Watson SR, Wilson KB, Zlotnik A;
XX WPI; 2003-468649/44.
XX P-PSDB; ADN39064.
XX Determining the presence or absence of a pathological cell in a patient,
XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.
XX Claim 8; SEQ ID NO 381; 1385pp; English.
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
XX whose expression is upregulated or downregulated in specific cancers or
XX other diseases such as angiogenic or fibrotic disorders, and to methods
XX of determining the presence or absence of a pathological cell in a
XX patient by detecting a nucleic acid at least 80% identical to those of
XX the invention or by detecting a polypeptide of the invention. The
XX invention also relates to expression vectors and host cells comprising a
XX nucleic acid of the invention; antibodies which specifically bind a
XX polypeptide of the invention; use of such antibodies for drug targeting;
XX and methods of screening for modulators of activity or expression of the
XX polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX antibodies and methods are useful for diagnosing, prognosing and treating
XX cancer and other conditions such as psoriasis, ischaemia, heart disease,
XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX neovascularisation syndromes, scarring and uterine fibroids. They may
XX also be useful in wound healing and in contraception. The present
XX sequence represents a nucleic acid sequence of the invention.
XX Sequence 7941 BP; 2406 A; 1651 C; 1598 G; 2286 T; 0 U; 0 Other;
Query Match 40.9%; Score 1263.2; DB 11; Length 7941;
Best Local Similarity 99.8%; Pred. No. 3.1e-264;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CACATACGACGACGACGATCTCCTGATCTATACACTGGAGGATTAACAAACAAA 60
DB 1 CACATACGACGACGACGATCTCCTGATCTATACACTGGAGGATTAACAAACAAA 60
QY 61 CAAAAAACAATTCCTTCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 120
DB 61 CAAAAAACAATTCCTTCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 120
QY 121 CGGCGAGGGCGCGAGACCGTCTGGAATGCGAATCTTAAGCGTTTCCTCGTTGCATT 180
DB 121 CGGCGAGGGCGCGAGACCGTCTGGAATGCGAATCTTAAGCGTTTCCTCGTTGCATT 180
QY 181 CAGCTCCTCTGTTGGCGCTGGAATGGGGCTAATGATCTACAGACAACAGAGAAA 240

DB 181 CAGCTCCTCTGTTGGCGCTGGAATGGGGCTAATGATCTACAGACAACAGAGAAA 240
QY 241 CTTGTTGAAGAGATTTGGCTGCTCTATACAGAGACATGAATCAAAAAAATTTGGGGAAG 300
DB 241 CTTGTTGAAGAGATTTGGCTGCTCTATACAGAGACATGAATCAAAAAAATTTGGGGAAG 300
QY 301 AAATATCAACATCTAATAGCCCAAAACAATCTCTCAATATTTGATGAAGATCTTACA 360
DB 301 AAATATCAACATCTAATAGCCCAAAACAATCTCTCAATATTTGATGAAGATCTTACA 360
QY 361 CAAGTAAATGTGAATCTTAAGAAAATTAAATTTCAAGGTTGGGATTAACATCATTTGAA 420
DB 361 CAAGTAAATGTGAATCTTAAGAAAATTAAATTTCAAGGTTGGGATTAACATCATTTGAA 420
QY 421 AACACATTCATTAACACATCGGGAACACAGTGAATTAATCTCACTAATGACTACCGT 480
DB 421 AACACATTCATTAACACATCGGGAACACAGTGAATTAATCTCACTAATGACTACCGT 480
QY 481 GTCAAGCGAGGATTTCAAGAAATGTTTAAAGCAAGCAAGATAAATTTTCACTGGGGA 540
DB 481 GTCAAGCGAGGATTTCAAGAAATGTTTAAAGCAAGCAAGATAAATTTTCACTGGGGA 540
QY 541 AAATGCAATATGTCATCTGATGATCGAGCATAGTTTAGAAGGACAAAAATTTCCACTT 600
DB 541 AAATGCAATATGTCATCTGATGATCGAGCATAGTTTAGAAGGACAAAAATTTCCACTT 600
QY 601 GAGATGCAAAATCTACTGCTTGTGATCGGACCGATTTTCAAGTTTGGAGGAAGCATCAAA 660
DB 601 GAGATGCAAAATCTACTGCTTGTGATCGGACCGATTTTCAAGTTTGGAGGAAGCATCAAA 660
QY 661 GGAAGAGGAGTTAAGAGCTTTTATCCATTTTGTGAGGTTGGGACAGAGAAATTTG 720
DB 661 GGAAGAGGAGTTAAGAGCTTTTATCCATTTTGTGAGGTTGGGACAGAGAAATTTG 720
QY 721 GATTTCAAAAGCGAATTTGATGAGTGCAGAAAGTGTAGTCTGTTTGGGAAGCAGCTGCT 780
DB 721 GATTTCAAAAGCGAATTTGATGAGTGCAGAAAGTGTAGTCTGTTTGGGAAGCAGCTGCT 780
QY 781 TTGATGCAATCTACTGCTTGTGAACTTCTGCAAACTCACTGACAGATTTACATTTAC 840
DB 781 TTGATGCAATCTACTGCTTGTGAACTTCTGCAAACTCACTGACAGATTTACATTTAC 840
QY 841 AATGCTCATTTGACATCTCTCCCTGCACACACAGTTGACTGATTTGTTTAAAGAT 900
DB 841 AATGCTCATTTGACATCTCTCCCTGCACACACAGTTGACTGATTTGTTTAAAGAT 900
QY 901 ACAGTTAGCATCTCTGAAAGCAGTTGGCTGTTTGTGAAAGTTCTTACAAATGCAACAA 960
DB 901 ACAGTTAGCATCTCTGAAAGCAGTTGGCTGTTTGTGAAAGTTCTTACAAATGCAACAA 960
QY 961 TCTGTTATGTCATGCTGATGAGTACTTACAAACAAATTTTTCGAGAGCAACAGTACAG 1020
DB 961 TCTGTTATGTCATGCTGATGAGTACTTACAAACAAATTTTTCGAGAGCAACAGTACAG 1020
QY 1021 TTCTCTAGACAGGTTTCTCTCATACACTGGAAGGAGAGATTCATGAAGCAGTTTGT 1080
DB 1021 TTCTCTAGACAGGTTTCTCTCATACACTGGAAGGAGAGATTCATGAAGCAGTTTGT 1080
QY 1081 AGTTTCAGAACCAAGAAATGTTTCAGGCTGACCCAGAGAAATATATACAGCCTTCTTGT 1140
DB 1081 AGTTTCAGAACCAAGAAATGTTTCAGGCTGACCCAGAGAAATATATACAGCCTTCTTGT 1140
QY 1141 TGGGAAGACCTCGAGTGGTTTATGATACCAATGATGAGAAAGTTTGAGAGTTTGACAG 1200
DB 1141 TGGGAAGACCTCGAGTGGTTTATGATACCAATGATGAGAAAGTTTGAGAGTTTGACAG 1200
QY 1201 CAGTTGATGAGAGGACCAACCAAGCATGAATTTTTCAGAGTGGCTATCAAGACTTG 1260
DB 1201 CAGTTGATGAGAGGACCAACCAAGCATGAATTTTTCAGAGTGGCTATCAAGACTTG 1260
QY 1261 GTAACATAT 1268

Dn	1261 GGTGCTAT 1268	Query Match	40.9%;	Score 1263.2;	DB 12;	Length 7941;
ADJ75139		Best Local Similarity	99.8%;	Pred. No. 3.1e-264;		
XX	ADJ75139 standard; DNA; 7941 BP.	Matches 1265;	Conservative	0;	Mismatches 3;	Indels 0;
XX	ADJ75139;	Gaps	0;			
XX						
DT	20-MAY-2004 (first entry)					
XX						
DE	Marker gene SEQ ID NO:391.					
XX						
KW	bronchial asthma; chronic obstructive pulmonary disease;					
KW	respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;					
KW	gene therapy; marker gene; gene; ds.					
XX						
OS	Homo sapiens.					
XX						
PN	EPI394274-A2.					
XX						
PD	03-MAR-2004.					
XX						
PF	04-AUG-2003; 2003EP-00254857.					
XX						
PR	06-AUG-2002; 2002JP-00229312.					
PR	20-MAR-2003; 2003JP-00077212.					
XX						
PA	(GENO-) GENOX RES INC.					
XX						
PI	Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuwara K;					
XX						
DR	WPI; 2004-193155/19.					
XX						
PT	Testing for bronchial asthma or chronic obstructive pulmonary disease by					
PT	comparing the expression level of a marker gene in a biological sample					
PT	from a subject with the expression level of the gene in a sample from a					
PT	healthy subject.					
XX						
PS	Claim 1; SEQ ID NO 391; 241pp; English.					
XX						
CC	The present invention describes a method of testing for bronchial asthma					
CC	or chronic obstructive pulmonary disease. The method comprises					
CC	determining the expression level of a marker gene in a biological sample					
CC	from a subject, comparing the expression level determined with the					
CC	expression level of the marker gene in a biological sample from a healthy					
CC	subject, and judging whether the subject has bronchial asthma or chronic					
CC	obstructive pulmonary disease. The marker gene comprises: (a) a group of					
CC	genes (S1) whose expression levels increase when respiratory epithelial					
CC	cells are stimulated with interleukin-13; or (b) a group of genes (S2)					
CC	whose expression levels decrease when respiratory epithelial cells are					
CC	stimulated with interleukin-13. Also described: (1) a reagent (1) for					
CC	testing for bronchial asthma or chronic obstructive pulmonary disease;					
CC	(2) a kit for screening for a candidate compound for a therapeutic agent					
CC	to treat bronchial asthma or chronic obstructive pulmonary disease; (3)					
CC	an animal model for bronchial asthma or chronic obstructive pulmonary					
CC	disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a					
CC	method for producing an animal model for bronchial asthma or chronic					
CC	obstructive pulmonary disease; (6) a therapeutic agent for bronchial					
CC	asthma or chronic obstructive pulmonary disease, comprising the compound,					
CC	a marker gene or an antisense nucleic acid corresponding to a portion of					
CC	the marker gene, a ribozyme, a polynucleotide that suppresses the					
CC	expression of the gene through an RNAi effect or an antibody recognising					
CC	a protein encoded by a marker gene; and (7) a DNA chip for testing for					
CC	bronchial asthma or a chronic obstructive pulmonary disease, on which a					
CC	probe has been immobilised to assay a marker gene. (1) has respiratory					
CC	and antiasthmatic activities, and can be used in gene therapy. The method					
CC	is useful for testing for or screening for a therapeutic agent for					
CC	bronchial asthma or chronic obstructive pulmonary disease. The present					
CC	sequence is used in the exemplification of the present invention.					
XX						
SQ	Sequence 7941 BP; 2406 A; 1651 C; 1598 G; 2286 T; 0 U; 0 Other;					

Db 1021 TTCTCTAGACAGGTGTTTCTCTCATACACTGGAAGAGAGATTTCATGACGAGTTTCT 1080
Qy 1081 AGTTTCAGAACCCAGAAATGTTTCAGGCTGACCCAGAGAAATATACACAGCTTCTTGTACA 1140
Db 1081 AGTTTCAGAACCCAGAAATGTTTCAGGCTGACCCAGAGAAATATACACAGCTTCTTGTACA 1140
Qy 1141 TGGGAAGACCTCGAGTCGTTTATGATACCATGATTGAGAAGTTTGACGTTTGTACCAAG 1200
Db 1141 TGGGAAGACCTCGAGTCGTTTATGATACCATGATTGAGAAGTTTGACGTTTGTACCAAG 1200
Qy 1201 CAGTTGGATGGAGAGACCAACCAAGCATGAATTTTTCAGACGATGGCTATCAAGACTTG 1260
Db 1201 CAGTTGGATGGAGAGACCAACCAAGCATGAATTTTTCAGACGATGGCTATCAAGACTTG 1260
Qy 1261 GTAACATAT 1268
Db 1261 GGTGCTAT 1268

RESULT 12
ABX76217
ID ABX76217 standard; DNA; 7944 BP.
XX
AC ABX76217;
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polynucleotide #86.
XX

Lung cancer-associated polynucleotide #86.

Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
ant inflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

Unidentified.

WO200286443-A2.

31-OCT-2002.

18-APR-2002; 2002WO-US012476.

18-APR-2001; 2001US-0284770P.

10-MAY-2001; 2001US-0290492P.

09-NOV-2001; 2001US-033245P.

13-NOV-2001; 2001US-0350666P.

29-NOV-2001; 2001US-0334370P.

12-APR-2002; 2002US-0372246P.

(BOSB-) EOS BIOTECHNOLOGY INC.

Pi Aziz N, Murray R;

WPI; 2003-093161/08.

P-PSDB; ABUS6493.

Detecting a lung cancer-associated transcript in a cell from a patient

for treating lung cancer, by contacting a biological sample from the
patient with a polynucleotide that exhibits increased or decreased
expression in lung cancer.

Claim 22; Page 255-257; 453pp; English.

The invention relates to a method for detecting a lung cancer-associated
transcript in a cell from a patient, comprising contacting a biological
sample from the patient with a polynucleotide that selectively hybridizes
to a sequence that is at least 80 % identical to a gene that exhibits
increased or decreased expression in lung cancer samples. Lung cancer-
associated polynucleotides and polypeptides are used for identifying a
compound that modulates a lung cancer-associated polypeptide, for
inhibiting proliferation of a lung cancer-associated cell to treat lung

CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC invention

XX Sequence 7944 BP; 2406 A; 1651 C; 1600 G; 2287 T; 0 U; 0 Other;

Qy Query Match 40.9%; Score 1263.2; DB 8; Length 7944;

Best Local Similarity 99.8%; Pred. No. 3.1e-264;

Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CACACATACGACGACGATCTCTTCGATCTATACACTGGAGGATTAAAAACAACAAA 60
Db 1 CACACATACGACGACGATCTCTTCGATCTATACACTGGAGGATTAAAAACAACAAA 60
Qy 61 CAAAAAAACATTTCTTCGCTCCCTCCCTCTCCACTCTGAGAGCAGAGCGCGCA 120
Db 61 CAAAAAAACATTTCTTCGCTCCCTCCCTCTCCACTCTGAGAGCAGAGCGCGCA 120
Qy 121 CGGCGAGGCGCGCAGACCGTCTGGAATGGAATCTTAAAGCGTTTCTCTGCTTGCATT 180
Db 121 CGGCGAGGCGCGCAGACCGTCTGGAATGGAATCTTAAAGCGTTTCTCTGCTTGCATT 180
Qy 181 CAGCTCTCTGTGTGTTGCGCGCTGGATTGGGCTAATGGATACTACAGACAACAGAGAAA 240
Db 181 CAGCTCTCTGTGTGTTGCGCGCTGGATTGGGCTAATGGATACTACAGACAACAGAGAAA 240
Qy 241 CTGTTGAAGAGATTGGCTGCTCTATACAGAGCACTGAATCAAAAAAATTTGGGGAAG 300
Db 241 CTGTTGAAGAGATTGGCTGCTCTATACAGAGCACTGAATCAAAAAAATTTGGGGAAG 300
Qy 301 AAATATCCAAACATGTAATAGCCCAAAACAATCTCTATCAATATTTGATGAGATCTTACA 360
Db 301 AAATATCCAAACATGTAATAGCCCAAAACAATCTCTATCAATATTTGATGAGATCTTACA 360
Qy 361 CAAGTAAATGTGAATCTTAAAGAACTTAAATTTCAAGGTTGGGATAAAACATCATTTGAA 420
Db 361 CAAGTAAATGTGAATCTTAAAGAACTTAAATTTCAAGGTTGGGATAAAACATCATTTGAA 420
Qy 421 AACACATTCATTCATAACACTGGGAAAAACAGTGGAAAAATTAATCTCACTAATGACTACCGT 480
Db 421 AACACATTCATTCATAACACTGGGAAAAACAGTGGAAAAATTAATCTCACTAATGACTACCGT 480
Qy 481 CTCAGCGGAGGAGTTTCAGAAATGTTTAAAGCAAGCAAGATAACTTTTCACCTGGGGA 540
Db 481 CTCAGCGGAGGAGTTTCAGAAATGTTTAAAGCAAGCAAGATAACTTTTCACCTGGGGA 540
Qy 541 AAATGCAATATGTCTATCTGATGGATCAGAGCATAGTTTAGAAGGACAAAAATTTCCACTT 600
Db 541 AAATGCAATATGTCTATCTGATGGATCAGAGCATAGTTTAGAAGGACAAAAATTTCCACTT 600
Qy 601 GAGATGCAAAATCTACTGTTTGAATGCGGACCGATTTTCAAGTTTGGAGAGCAGTCAAA 660
Db 601 GAGATGCAAAATCTACTGTTTGAATGCGGACCGATTTTCAAGTTTGGAGAGCAGTCAAA 660
Qy 661 GGAAGGGAAGTTTAAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAATTTG 720
Db 661 GGAAGGGAAGTTTAAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAATTTG 720
Qy 721 GATTTCAAGAGGATTTATGATGGAGTCGAAAGTGTAGTCTGTTTGGAGAGCAGCTGCT 780
Db 721 GATTTCAAGAGGATTTATGATGGAGTCGAAAGTGTAGTCTGTTTGGAGAGCAGCTGCT 780
Qy 781 TTAGATCAATTCATCTGTTGAACTCTTCGCAAACTCAACTGACAGATTTTACATTTAC 840
Db 781 TTAGATCAATTCATCTGTTGAACTCTTCGCAAACTCAACTGACAGATTTTACATTTAC 840

QY 841 AATGGCTCATGACATCTCCTCCCTGCACACACAGTTGACTGGATTGTTTTAAAGAT 900
Db |||||
841 AATGGCTCATGACATCTCCTCCCTGCACACACAGTTGACTGGATTGTTTTAAAGAT 900
QY 901 ACAGTTAGCATCTCTGAAGCCAGTTGGCTGTTTTTTGTGAAGTTCTTACAAATGCAACAA 960
Db |||||
901 ACAGTTAGCATCTCTGAAGCCAGTTGGCTGTTTTTTGTGAAGTTCTTACAAATGCAACAA 960
QY 961 TCTGTTATGTATGCTGATGGACTACTTACAAAACAATTTTCGAGAGCAACAGTACAAAG 1020
Db |||||
961 TCTGTTATGTATGCTGATGGACTACTTACAAAACAATTTTCGAGAGCAACAGTACAAAG 1020
QY 1021 TTCTCTAGACAGTGTGTTTCTCTCATACACTGAAAGGAAGATTTCATGAAGCAGTTTGT 1080
Db |||||
1021 TTCTCTAGACAGTGTGTTTCTCTCATACACTGAAAGGAAGATTTCATGAAGCAGTTTGT 1080
QY 1081 AGTTCAGAACACAGAAATGTTTCAGGCTGACCCACAGAGATTATACCAGCTTCTTGTGTACA 1140
Db |||||
1081 AGTTCAGAACACAGAAATGTTTCAGGCTGACCCACAGAGATTATACCAGCTTCTTGTGTACA 1140
QY 1141 TGGAAAGACCTCGAGTCGTTTATGATACCATGATTGAGAAGTTTGTGCAAGTTTGTACCAAG 1200
Db |||||
1141 TGGAAAGACCTCGAGTCGTTTATGATACCATGATTGAGAAGTTTGTGCAAGTTTGTACCAAG 1200
QY 1201 CAGTTGGATGGAGGACCAACCAAGCATGAATTTTTCAGAGATGGCTATCAAGACTTG 1260
Db |||||
1201 CAGTTGGATGGAGGACCAACCAAGCATGAATTTTTCAGAGATGGCTATCAAGACTTG 1260
QY 1261 GTAACAT 1268
Db |||||
1261 GGTGCTAT 1268

RESULT 13
ABX76416
ID ABX76416 standard; DNA; 7944 BP.
XX
AC ABX76416;
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polynucleotide #280.
XX
KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW antiinflammatory; antiaethmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
XX WO200286443-A2.
PN
XX
XX 31-OCT-2002.
PD
XX
PF 18-APR-2002; 2002WO-US012476.
XX
XX 18-APR-2001; 2001US-0284770P.
PR
PR 09-MAY-2001; 2001US-0290492P.
PR
PR 13-NOV-2001; 2001US-0339245P.
PR
PR 29-NOV-2001; 2001US-0350666P.
PR
PR 29-NOV-2001; 2001US-0334370P.
PR
PR 12-APR-2002; 2002US-0372246P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
PA
XX
XX Aziz N, Murray R;
PI
XX
XX WPI; 2003-093161/08.
DR
DR P-PSDB; ABU56687.
XX
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient

for treating lung cancer, by contacting a biological sample from the patient with a polynucleotide that exhibits increased or decreased expression in lung cancer.

Claim 22; Page 402-403; 453pp; English.

The invention relates to a method for detecting a lung cancer-associated transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridises to a sequence that is at least 80 % identical to a gene that exhibits increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the invention

Sequence 7944 BP; 2406 A; 1651 C; 1600 G; 2287 T; 0 U; 0 Other;
Query Match 40.9%; Score 1263.2; DB 8; Length 7944;
Best Local Similarity 99.8%; Pred. No. 3.1e-264;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACACATACGACGACGATCTGATTCGATCTATACACTGGAGATTAAACAACAAA 60
Db |||||
1 CACACATACGACGACGATCTGATTCGATCTATACACTGGAGATTAAACAACAAA 60
QY 61 CAAAAAAAACATTTCTTGGCTCCCTCCCTCTCCACTCTGAGAAGCAGAGAGCCGCA 120
Db |||||
61 CAAAAAAAACATTTCTTGGCTCCCTCCCTCTCCACTCTGAGAAGCAGAGAGCCGCA 120
QY 121 CGGCGAGGGGCCGAGACCGTCTGGAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180
Db |||||
121 CGGCGAGGGGCCGAGACCGTCTGGAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180
QY 181 CAGCTCTCTGTGTTTGGCCGCTGGATTGGCTAATGGATCTACACACACAGAAAA 240
Db |||||
181 CAGCTCTCTGTGTTTGGCCGCTGGATTGGGCTAATGGATCTACACACACAGAAAA 240
QY 241 CTCTGTGAAGAGATTGGCTGCTCTATACAGAGACACTGAATCAAAAAAATTTGGGGAAAG 300
Db |||||
241 CTCTGTGAAGAGATTGGCTGCTCTATACAGAGACACTGAATCAAAAAAATTTGGGGAAAG 300
QY 301 AAATATCCAAACATGTAATAGCCCCAAAAACAATCTCTATCAATATGTAGTGAAGATCTTACA 360
Db |||||
301 AAATATCCAAACATGTAATAGCCCCAAAAACAATCTCTATCAATATGTAGTGAAGATCTTACA 360
QY 361 CAAGTAAATGTGAATCTTAAGAACTTAAATTTTCAAGGTTGGGATTAACATCATTTGAA 420
Db |||||
361 CAAGTAAATGTGAATCTTAAGAACTTAAATTTTCAAGGTTGGGATTAACATCATTTGAA 420
QY 421 AACACATTCATTTCATACACCTGGGAAAAACAGTGGAAAAATTAATCTCACTAATGACTACCGT 480
Db |||||
421 AACACATTCATTTCATACACCTGGGAAAAACAGTGGAAAAATTAATCTCACTAATGACTACCGT 480
QY 481 GTCAGCGGAGGAGTTTCAGAAATGGTGTTHAAGCAAGCAAGATAAATTTTTCACCTGGGGA 540
Db |||||
481 GTCAGCGGAGGAGTTTCAGAAATGGTGTTHAAGCAAGCAAGATAAATTTTTCACCTGGGGA 540
QY 541 AAATGCAATATGTCATCTGATGGATCAGACATAGTTTAGAAGGACCAAAATTTTCCACTT 600
Db |||||
541 AAATGCAATATGTCATCTGATGGATCAGACATAGTTTAGAAGGACCAAAATTTTCCACTT 600
QY 601 GAGATGCAAAATCTACTGCTTTGATGCGGACCGGATTTTCAAGTTTGTAGGAAAGCAGTCAAA 660

```
Db 601 GAGATCAAACTACTGCTTTGATGCGACCGATTTTCAAGTTTTCAGGAGCAGTCAAA 660
Qy 661 GGAAGAGGAGGTAAGAGCTTTATCCATTTGTTGAGGTTGGGACAGAAATTTG 720
Db 661 GGAAGAGGAGGTAAGAGCTTTATCCATTTGTTGAGGTTGGGACAGAAATTTG 720
Qy 721 GATTTCAAGCGGATTTATGATGAGTTCGAAAGTGTGTAGTTCGTTTGGGAGCAGGCTGCT 780
Db 721 GATTTCAAGCGGATTTATGATGAGTTCGAAAGTGTGTAGTTCGTTTGGGAGCAGGCTGCT 780
Qy 781 TTAGATCCATTCATCTGTTGAACTTTCGCCAACTCAACTGCAAGTATTACATTAC 840
Db 781 TTAGATCCATTCATCTGTTGAACTTTCGCCAACTCAACTGCAAGTATTACATTAC 840
Qy 841 AATGGCTCATTCATCTCTCCCTGCACAGACACAGTTGACTGGATTGTTTAAAGAT 900
Db 841 AATGGCTCATTCATCTCTCCCTGCACAGACACAGTTGACTGGATTGTTTAAAGAT 900
Qy 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTGTGAAGTTCCTTACAATGCAACAA 960
Db 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTGTGAAGTTCCTTACAATGCAACAA 960
Qy 961 TCTGGTTATGTCATGTCGTGACTACTTACAAAACAAATTTTCGAGAGCAACAGTACAAG 1020
Db 961 TCTGGTTATGTCATGTCGTGACTACTTACAAAACAAATTTTCGAGAGCAACAGTACAAG 1020
Qy 1021 TTCTCTAGACAGTGTTCCTCATACACTGGAAGGAGAGATTCATGAAGCAGTTGT 1080
Db 1021 TTCTCTAGACAGTGTTCCTCATACACTGGAAGGAGAGATTCATGAAGCAGTTGT 1080
Qy 1081 AGTTCAAGAACAGAAATGTTTCAGGCTGACCCAGAGAAATATACAGCCTCTCTGTACA 1140
Db 1081 AGTTCAAGAACAGAAATGTTTCAGGCTGACCCAGAGAAATATACAGCCTCTCTGTACA 1140
Qy 1141 TGGGAAAGACCTCGAGTCGTTTATGATACCATGATGAGAAGTTTCAGATTGTTACCAAG 1200
Db 1141 TGGGAAAGACCTCGAGTCGTTTATGATACCATGATGAGAAGTTTCAGATTGTTACCAAG 1200
Qy 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGATGATGATGATGATGATGATGATGATGATG 1260
Db 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGATGATGATGATGATGATGATGATGATGATG 1260
Qy 1261 GTAACAT 1268
Db 1261 GGTGCTAT 1268
```

RESULT 14

ADN39725

ADN39725 standard; cDNA; 7944 BP.

AC ADN39725;

XX 17-JUN-2004 (first entry)

XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO: C97.

DE Human; differential expression; cancer; angiogenic disorder;

DE fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;

DE inflammatory disease; autoimmune disease;

DE retinal neovascularization syndrome; scarring; uterine fibroid;

DE detection; diagnosis; prognosis; drug screening; drug targeting;

DE wound healing; contraception; cytostatic; cardiac; immunomodulatory;

DE vulnerability; gene therapy; vaccine; gene; ss.

XX Homo sapiens.

XX WO2003042661-A2.

XX 22-MAY-2003.

PD 13-NOV-2002; 2002WO-US036810.

XX 13-NOV-2002; 2002WO-US036810.

```
XX 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368099P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX (BOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DR, Murray R, Watson SR, Wilson KE, Zlotnik A;
PI WPI; 2003-468649/44.
DR P-PSDB; ADN39942.
```

Determining the presence or absence of a pathological cell in a patient, useful for diagnosing, prognosing or treating cancer, comprises detecting a nucleic acid in a biological sample.

Claim 8; SEQ ID NO C97; 1395pp; English.

The invention relates to nucleic acids and proteins (ADN38683-ADN40064) whose expression is upregulated or downregulated in specific cancers or other diseases such as angiogenic or fibrotic disorders, and to methods of determining the presence or absence of a pathological cell in a patient by detecting a nucleic acid at least 80% identical to those of the invention or by detecting a polypeptide of the invention. The invention also relates to expression vectors and host cells comprising a nucleic acid of the invention; antibodies which specifically bind a polypeptide of the invention; use of such antibodies for drug targeting; and methods of screening for modulators of activity or expression of the polypeptides and nucleic acids. The nucleic acids, polypeptides, antibodies and methods are useful for diagnosing, prognosing and treating cancer and other conditions such as psoriasis, ischaemia, heart disease, atherosclerosis, inflammatory diseases, autoimmune diseases, retinal neovascularization syndromes, scarring and uterine fibroids. They may also be useful in wound healing and in contraception. The present sequence represents a nucleic acid sequence of the invention.

Sequence 7944 BP; 2406 A; 1651 C; 1600 G; 2287 T; 0 U; 0 Other;

Query Match 40.9%; Score 1263.2; DB 11; Length 7944;

Best Local Similarity 99.8%; Pred. No. 3.1e-264;

Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy 1 CACATACGCGACGACGATCTCACTTCGATCTATATACCTGGAGGATTAACAACAAA 60
Db 1 CACATACGCGACGACGATCTCACTTCGATCTATATACCTGGAGGATTAACAACAAA 60
Qy 61 CAAAAAACAATTTCTTCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 120
Db 61 CAAAAAACAATTTCTTCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 120
Qy 121 CGCGAGGGGCGCGACGCGCTCTGAAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180
Db 121 CGCGAGGGGCGCGACGCGCTCTGAAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180
Qy 181 CAGCTCTCTGTGTTTCCCGCTGGATTGGGCTAATGATATACTACAGACACAGAGAAA 240
Db 181 CAGCTCTCTGTGTTTCCCGCTGGATTGGGCTAATGATATACTACAGACACAGAGAAA 240
```

QY 241 CTTGTTGAAGAGATTGGCTGCTCTATACAGGAGCACTGAATCAAAAAAATTGGGAAAG 300
Db |
QY 241 CTTGTTGAAGAGATTGGCTGCTCTATACAGGAGCACTGAATCAAAAAAATTGGGAAAG 300
Db |
QY 301 AAATATCAACATGTAATAGCCCAAAACAATCTCTATCAATATTGATGAAGATCTTACA 360
Db |
QY 301 AAATATCAACATGTAATAGCCCAAAACAATCTCTATCAATATTGATGAAGATCTTACA 360
Db |
QY 361 CAAGTAATGTGAATCTTAAAGAACTTAAATTTCAAGGTTGGGATAAACAATCATTTGAA 420
Db |
QY 361 CAAGTAATGTGAATCTTAAAGAACTTAAATTTCAAGGTTGGGATAAACAATCATTTGAA 420
Db |
QY 421 AACACATTCATTCATCAAC 480
Db |
QY 421 AACACATTCATTCATCAAC 480
Db |
QY 481 GTCAGCGAGGAGTTTCAAGAAATGTTGTTAAAGCAAGCAAGATAAATTTTCACTGGGA 540
Db |
QY 481 GTCAGCGAGGAGTTTCAAGAAATGTTGTTAAAGCAAGCAAGATAAATTTTCACTGGGA 540
Db |
QY 541 AAATGCAATATGTCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db |
QY 541 AAATGCAATATGTCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db |
QY 601 GAGATGCAATCTACTGCTTTGATGCGGACCGATTTTCAAGTTTGGAGAGCAATCAAA 660
Db |
QY 601 GAGATGCAATCTACTGCTTTGATGCGGACCGATTTTCAAGTTTGGAGAGCAATCAAA 660
Db |
QY 661 GGAAAGGGAAGTTAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAATTTG 720
Db |
QY 661 GGAAAGGGAAGTTAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAATTTG 720
Db |
QY 721 GATTTCAAAGCGATTTATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 780
Db |
QY 721 GATTTCAAAGCGATTTATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 780
Db |
QY 781 TTAGTCCATTCATCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db |
QY 781 TTAGTCCATTCATCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db |
QY 841 AATGCTCATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db |
QY 841 AATGCTCATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db |
QY 901 ACAATTAGCATCTCTGAAGCCAGTTGGCTGTTTGTGAAGTCTTCAATGCAACAA 960
Db |
QY 901 ACAATTAGCATCTCTGAAGCCAGTTGGCTGTTTGTGAAGTCTTCAATGCAACAA 960
Db |
QY 961 TCTGTTATGTCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1020
Db |
QY 961 TCTGTTATGTCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1020
Db |
QY 1021 TTCTCTAGACAGGTTGTTTCTCTCATACATGCTGAAAGGAGAGATTCATGAAGCAGTTGT 1080
Db |
QY 1021 TTCTCTAGACAGGTTGTTTCTCTCATACATGCTGAAAGGAGAGATTCATGAAGCAGTTGT 1080
Db |
QY 1081 AGTTGAGAACAGAAATGTTTCAGGCTGACCCAGAGAAATATACAGGCTTCTTGTTCACA 1140
Db |
QY 1081 AGTTGAGAACAGAAATGTTTCAGGCTGACCCAGAGAAATATACAGGCTTCTTGTTCACA 1140
Db |
QY 1141 TGGGAAAGACCTCGAGTCTGTTATGATACCATGATGATGAGAGTTTGCAGTTTGTACCG 1200
Db |
QY 1141 TGGGAAAGACCTCGAGTCTGTTATGATACCATGATGATGAGAGTTTGTGAGTTTGTACCG 1200
Db |
QY 1201 CAGTTGGATGAGAGGACCAACCAAGCAATGAAATTTTGCAGATGGCTATCAAGACTTG 1260
Db |
QY 1261 GTAACTAT 1268
Db |
QY 1261 GTGCTAT 1268
Db |

RESULT 15
ADN39053
ID ADN39053 standard; cDNA; 7944 BP.
XX
AC ADN39053;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:371.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulneryary; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
FN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
(EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Gineburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI; 2003-468649/44.
DR P-PSDB; ADN39054.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 8; SEQ ID NO 371; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,

CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.

XX
SQ Sequence 7944 BP; 2406 A; 1651 C; 1600 G; 2287 T; 0 U; 0 Other;

Query Match 40.9%; Score 1263.2; DB 11; Length 7944;
Best Local Similarity 99.8%; Pred. No. 3.1e-264;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1	CACACATACGACGACGATCTC	901	ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTGTTTTCTGGAAGTTCTTACAATGCAACAA	960
DB	1	CACACATACGACGACGATCTC	961	TCTGGTTATGTCATGCTCATG	1020
QY	61	CAAAAAAATCTTCTCGCTCC	1021	TTCTCTAGACAGGTGTTTTCT	1080
DB	61	CAAAAAAATCTTCTCGCTCC	1081	AGTTTCAGAACCCAGAAAAAT	1140
QY	121	CGGCGAGGGGCGCAGACCG	1141	TGGGAAAGACCTCCGAGTCG	1200
DB	121	CGGCGAGGGGCGCAGACCG	1201	CAGTTGGATGGAGAGGACCA	1260
QY	181	CAGCTCCTCTGTGTTTGG	1261	GTAACAT 1268	
DB	181	CAGCTCCTCTGTGTTTGG	1268	GATGCTAT 1268	
QY	241	CTTGTGTAAGAGATTGGCT			
DB	241	CTTGTGTAAGAGATTGGCT			
QY	301	AAATATCCAAATGTAATAG			
DB	301	AAATATCCAAATGTAATAG			
QY	361	CAAGTAAATGTGAATCTT			
DB	361	CAAGTAAATGTGAATCTT			
QY	421	AAACATTTCAATTCATAC			
DB	421	AAACATTTCAATTCATAC			
QY	481	GTGAGCGGAGGAGTTTCA			
DB	481	GTGAGCGGAGGAGTTTCA			
QY	541	AAATGCAATATGTCTCAT			
DB	541	AAATGCAATATGTCTCAT			
QY	601	GAGATGCAATCTACTGCT			
DB	601	GAGATGCAATCTACTGCT			
QY	661	GGAAAGGGAAGTTAAGAG			
DB	661	GGAAAGGGAAGTTAAGAG			
QY	721	GATTTCAAAAGCGAATTA			
DB	721	GATTTCAAAAGCGAATTA			
QY	781	TTAGATCCCAATCTACTG			
DB	781	TTAGATCCCAATCTACTG			
QY	841	AATGGCTCAATGACATCT			
DB	841	AATGGCTCAATGACATCT			
QY	901	ACAGTTAGCATCTCTGAA			
DB	901	ACAGTTAGCATCTCTGAA			

Search completed: August 30, 2005, 12:11:44
Job time : 1589 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 11:45:24 ; Search time 1910 Seconds
(without alignments)
10588.843 Million cell updates/sec

Title: US-09-983-000A-1
Perfect score: 3091
Sequence: 1 cacacatgcacgcacgat.....ttatgtagtaagtgtatata 3091

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgm2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgm2_6/ptodata/2/pubpna/PT_NEW_PUB.seq.*
- 3: /cgm2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgm2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgm2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgm2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 7: /cgm2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgm2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgm2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgm2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgm2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgm2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgm2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgm2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgm2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgm2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgm2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgm2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgm2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgm2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgm2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgm2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 23: /cgm2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgm2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgm2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgm2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	3091	100.0	3091	10	US-09-983-000A-1
2	3091	100.0	3091	21	US-10-652-981-1
3	1263.2	40.9	5367	17	US-10-295-027-373
4	1263.2	40.9	5367	17	US-10-295-027-1044
5	1263.2	40.9	7941	9	US-09-954-456-1812
6	1263.2	40.9	7941	9	US-09-816-703A-1
7	1263.2	40.9	7941	10	US-09-983-000A-5

8	1263.2	40.9	7941	17	US-10-295-027-381
9	1263.2	40.9	7941	21	US-10-843-641A-4839
10	1263.2	40.9	7941	21	US-10-652-981-5
11	1263.2	40.9	7941	21	US-10-936-626-32
12	1263.2	40.9	7941	21	US-10-938-061-32
13	1263.2	40.9	7941	21	US-10-295-027-371
14	1263.2	40.9	7941	17	US-10-295-027-1043
15	1263.2	40.9	7941	21	US-10-852-335A-6
16	1263.2	40.9	8058	10	US-09-983-000A-3
17	1263.2	40.9	8058	21	US-10-652-981-3
18	1263.2	40.9	8292	20	US-10-723-860-5522
19	1261.6	40.8	5343	17	US-10-295-027-375
20	1261.6	40.8	5481	17	US-10-295-027-379
21	1251.2	40.5	5363	17	US-10-295-027-377
22	1251.2	40.5	5363	17	US-10-295-027-1045
23	1250.6	40.5	5481	17	US-10-295-027-1046
24	1242.2	40.2	3350	20	US-10-723-860-950
25	1116.2	36.1	3340	17	US-10-295-027-1047
26	1116.2	36.1	4220	17	US-10-295-027-1049
27	1114.6	36.1	4347	17	US-10-295-027-1042
28	1113	36.0	4480	17	US-10-295-027-1048
29	935.2	30.3	6887	22	US-10-698-190-31
30	926.4	30.0	6801	15	US-10-205-219-120
31	926.4	30.0	6801	22	US-10-698-190-29
32	204.6	6.6	774	13	US-10-027-632-131144
33	204.6	6.6	774	13	US-10-027-632-161783
34	204.6	6.6	774	17	US-10-027-632-131144
35	204.6	6.6	774	17	US-10-027-632-161783
36	189	6.1	5787	21	US-10-887-553A-434
37	189	6.1	6687	20	US-10-723-860-5795
38	60.2	1.9	18817	15	US-10-240-485-47
39	58.4	1.9	6465	15	US-10-311-455-958
40	58.4	1.9	8056	20	US-10-473-126-386
41	58	1.9	3673778	16	US-10-312-841-1
42	57.4	1.9	7498	15	US-10-311-455-230
43	54.4	1.8	38678	22	US-10-893-315-136
44	54.4	1.8	38684	22	US-10-893-315-154
45	54.2	1.8	8056	20	US-10-473-126-386

ALIGNMENTS

RESULT 1

US-09-983-000A-1
; Sequence 1, Application US/09983000A
; Publication No. US20030118585A1
; GENERAL INFORMATION:
; APPLICANT: AGY Therapeutics
; APPLICANT: Melcher, Thorsten
; APPLICANT: Mueller, Sabine
; APPLICANT: Chin, Daniel
; TITLE OF INVENTION: USE OF PROTEIN BIOMOLECULAR TARGETS IN THE TREATMENT AND VISUALIZ
; FILE OF INVENTION: OF BRAIN TUMORS
; FILE REFERENCE: 263/180 -- Peagleman -- AGY
; CURRENT APPLICATION NUMBER: US/09/983,000A
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 3091
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: gene
; LOCATION: (1)..(3091)
; OTHER INFORMATION: PTP-zeta SM1 exon 9 variant
; NAME/KEY: VARSPLIC
; LOCATION: (1262)..(1262)
; OTHER INFORMATION: Alternative splice site
; NAME/KEY: misc feature
; LOCATION: (1273)..(3091)
; OTHER INFORMATION: 3' Untranslated Region

; NAME/KEY: CDS
; LOCATION: (148)..(1272)
; OTHER INFORMATION:
US-09-983-000a-1

Query Match
Best Local Similarity 100.0%; Score 3091; DB 10; Length 3091;
Matches 3091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CACACATACGACGACGATCTCACTTCGATCTATACACTGGAGGATTAACAAACAAA	60
Db	1	CACACATACGACGACGATCTCACTTCGATCTATACACTGGAGGATTAACAAACAAA	60
Qy	61	CAAAAAAATTCCTTCGCTCCCTCCCTCTCCACTCTGAGAGCAGAGGCCGCA	120
Db	61	CAAAAAAATTCCTTCGCTCCCTCCCTCTCCACTCTGAGAGCAGAGGCCGCA	120
Qy	121	CGCGAGGGCGCAGACCGCTCGGAATGCAATCCTTAAAGCGTTTCCTCGCTTGCA	180
Db	121	CGCGAGGGCGCAGACCGCTCGGAATGCAATCCTTAAAGCGTTTCCTCGCTTGCA	180
Qy	181	CAGCTCCTCTGTTTGGCGCTGGATTTGGCTAATGGATCTACAGACAAACAGAGAAA	240
Db	181	CAGCTCCTCTGTTTGGCGCTGGATTTGGCTAATGGATCTACAGACAAACAGAGAAA	240
Qy	241	CTTGTGAAGAGATTGGCTGGCTCTATACAGAGCAGCTGAATCAAAAAATTTGGGAA	300
Db	241	CTTGTGAAGAGATTGGCTGGCTCTATACAGAGCAGCTGAATCAAAAAATTTGGGAA	300
Qy	301	AAATATCAACATGTAATAGCCCAAAACAACTCTCTATCAATATTTGATGAAGATCTTACA	360
Db	301	AAATATCAACATGTAATAGCCCAAAACAACTCTCTATCAATATTTGATGAAGATCTTACA	360
Qy	361	CAAGTAATGTGAATCTTTAAGAACTTAAATTTCAAGGTTGGGATAAACAATCAATGGAA	420
Db	361	CAAGTAATGTGAATCTTTAAGAACTTAAATTTCAAGGTTGGGATAAACAATCAATGGAA	420
Qy	421	AACACATTCATTAACACACTGGGAAACAGTGGAAATTAATCTCACTAATGACTACCGT	480
Db	421	AACACATTCATTAACACACTGGGAAACAGTGGAAATTAATCTCACTAATGACTACCGT	480
Qy	481	GTACGCGAGGAGTTTCAAGAAATGGTTTAAAGCAAGCAAGATACTTTTCACTGGGGA	540
Db	481	GTACGCGAGGAGTTTCAAGAAATGGTTTAAAGCAAGCAAGATACTTTTCACTGGGGA	540
Qy	541	AAATGCAATATGTCTATCTGATGGATCAGAGCATAGTTTGAAGGACAAAATTTCCACTT	600
Db	541	AAATGCAATATGTCTATCTGATGGATCAGAGCATAGTTTGAAGGACAAAATTTCCACTT	600
Qy	601	GAGATGCAATCTACTGCTTTGATGCGGACCGATTTTCAAGTTTGAAGGAGCAGTCAAA	660
Db	601	GAGATGCAATCTACTGCTTTGATGCGGACCGATTTTCAAGTTTGAAGGAGCAGTCAAA	660
Qy	661	GGAAAGGGAAGTTAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAAATTTG	720
Db	661	GGAAAGGGAAGTTAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAAATTTG	720
Qy	721	GATTTCAAAGCGATTTAATGAGGAGTCGAAAGTGTAGTCTGTTTGGGAGCAGGCTGCT	780
Db	721	GATTTCAAAGCGATTTAATGAGGAGTCGAAAGTGTAGTCTGTTTGGGAGCAGGCTGCT	780
Qy	781	TTAGATCCATTCATCTCTGACCTCTGCCAACTCTCAACTCAACTGACAGTATTTACATTTAC	840
Db	781	TTAGATCCATTCATCTCTGACCTCTGCCAACTCTCAACTCAACTGACAGTATTTACATTTAC	840
Qy	841	AATGGCTCATTTGACATCTCTCCCTGACAGACAGTTGACTGATTTGTTTAAAGAT	900
Db	841	AATGGCTCATTTGACATCTCTCCCTGACAGACAGTTGACTGATTTGTTTAAAGAT	900
Qy	901	ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTGTGAAGTTCTTACAAATGCAACA	960
Db	901	ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTGTGAAGTTCTTACAAATGCAACA	960

Qy	961	TCTGGTTATGTCATGCTGATGGACTCTTACAAAAAATTTTCGAGAGCAACAGTACAAG	1020
Db	961	TCTGGTTATGTCATGCTGATGGACTCTTACAAAAAATTTTCGAGAGCAACAGTACAAG	1020
Qy	1021	TTCTCTAGACAGGTTGTTTCTCTCATACACTGGAAAGGAAGAGATTCATGAAGCAGTTTGT	1080
Db	1021	TTCTCTAGACAGGTTGTTTCTCTCATACACTGGAAAGGAAGAGATTCATGAAGCAGTTTGT	1080
Qy	1081	AGTTTCAGAACAGAAAATGTTTCAGGCTGACCCAGAGAAATTTATACCAGCTTCTTGTTACA	1140
Db	1081	AGTTTCAGAACAGAAAATGTTTCAGGCTGACCCAGAGAAATTTATACCAGCTTCTTGTTACA	1140
Qy	1141	TGGGAAGAGCCTCGAGTCGTTTATGATACCATGATTCGAGAAAGTTTGAGTTTGTACCA	1200
Db	1141	TGGGAAGAGCCTCGAGTCGTTTATGATACCATGATTCGAGAAAGTTTGAGTTTGTACCA	1200
Qy	1201	CAGTTGATGGAGAGGACCAAAACCAAGCATGAATTTTTCAGAGTGCTATCAAGACTTG	1260
Db	1201	CAGTTGATGGAGAGGACCAAAACCAAGCATGAATTTTTCAGAGTGCTATCAAGACTTG	1260
Qy	1261	GTAATATATGATCAGTTGTTTTCATAGAGGTAAATTAATTTTCCAAAGGTAAG	1320
Db	1261	GTAATATATGATCAGTTGTTTTCATAGAGGTAAATTAATTTTCCAAAGGTAAG	1320
Qy	1321	AACCTTCAAAATGTTGATATATTTTCTCCATTACTTTTTCAGACTTTATGTGAAGTGG	1380
Db	1321	AACCTTCAAAATGTTGATATATTTTCTCCATTACTTTTTCAGACTTTATGTGAAGTGG	1380
Qy	1381	GGTAGGCTGAGTATTTTAAATTTTAAAAAAAATTTTAAATTTAGAGCTATACATAAATTA	1440
Db	1381	GGTAGGCTGAGTATTTTAAATTTTAAAAAAAATTTTAAATTTAGAGCTATACATAAATTA	1440
Qy	1441	TGTTTAAAGTTACATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA	1500
Db	1441	TGTTTAAAGTTACATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA	1500
Qy	1501	GTAGATACATATGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA	1560
Db	1501	GTAGATACATATGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA	1560
Qy	1561	TTGTAGAAATAGTGCAGAAATCTCAGTGCCCTGATACATTTTATATTTGTTGCTTCCATTAC	1620
Db	1561	TTGTAGAAATAGTGCAGAAATCTCAGTGCCCTGATACATTTTATATTTGTTGCTTCCATTAC	1620
Qy	1621	GCTATATCAGCAGGAAAAGTAGAGTAGGGGACATACAGTCTCTTTTGTGACACAAA	1680
Db	1621	GCTATATCAGCAGGAAAAGTAGAGTAGGGGACATACAGTCTCTTTTGTGACACAAA	1680
Qy	1681	AAATTTTCAGATAACAGCTGGAAAGTCATGATGGGTGAGAACTTTTGGGGATGTAAGAAA	1740
Db	1681	AAATTTTCAGATAACAGCTGGAAAGTCATGATGGGTGAGAACTTTTGGGGATGTAAGAAA	1740
Qy	1741	ACATTTCTTAACAAAAGATCCACCCCTGCTCCCTCCACAGGCGCATGCGAATAAAGTAC	1800
Db	1741	ACATTTCTTAACAAAAGATCCACCCCTGCTCCCTCCACAGGCGCATGCGAATAAAGTAC	1800
Qy	1801	AGATTCCTTTGCTGCTGAGCATGTCAGTATTTAACTTTTCTCTGCTGCTGAGGAGTGTG	1860
Db	1801	AGATTCCTTTGCTGCTGAGCATGTCAGTATTTAACTTTTCTCTGCTGCTGAGGAGTGTG	1860
Qy	1861	GCCATAGATTAGGGTGTAGTTTGAACAACTTTCATCTGGATGTAGGTCCAGAAAGTCCCCA	1920
Db	1861	GCCATAGATTAGGGTGTAGTTTGAACAACTTTCATCTGGATGTAGGTCCAGAAAGTCCCCA	1920
Qy	1921	CTGAGGTTTAAAGGACACTGCACTCTGCACTCAGGCACTTAGAGTCTCTGCAAGTCTGGG	1980
Db	1921	CTGAGGTTTAAAGGACACTGCACTCTGCACTCAGGCACTTAGAGTCTCTGCAAGTCTGGG	1980
Qy	1981	AACCTGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2040
Db	1981	AACCTGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2040
Qy	2041	ATAAAAATTTTAGTAAACCTTTTAAATTCAGTTGCTGGAATATGGAGACACAAATGACCTGGG	2100

2041 ATAAAAATTTAGTAACCTTTTAATTCAGTTGCGTGGAAATATGAGACACAATGACCTGGG 2100
2101 AAAATCGTCAAAATAAGTAAATGTTTATTTCAATAATTACGTGAAGAAGTAAT 2160
2101 AAAATCGTGAATTAATAGTAATAAATGTTTATTTCAATAATTACGTGAAGAAGTAAT 2160
2161 TCTATTACTGTTCTTGGCATATATTTGTCAAGAAAAGAGATAACTTAGTTGTTCACTTT 2220
2161 TCTATTACTGTTCTTGGCATATATTTGTCAAGAAAAGAGATAACTTAGTTGTTCACTTT 2220
2221 TTCACATTTGCTCTGTTTGGAAATGCCCCCAATTTATTTGCTAAATAATTAATTTTAA 2280
2221 TTCACATTTGCTCTGTTTGGAAATGCCCCCAATTTATTTGCTAAATAATTAATTTTAA 2280
2281 GTTTGTAGTACTAAATTTATGAATTTGATGAGTTCTGGCTTAAATGAAACTTCTTGAAC 2340
2281 GTTTGTAGTACTAAATTTATGAATTTGATGAGTTCTGGCTTAAATGAAACTTCTTGAAC 2340
2341 TAAATCTGATTTTAAAAAGCAAAAAAAGGCTAGCTTTCCAGTTCTTTCATAA 2400
2341 TAAATCTGATTTTAAAAAGCAAAAAAAGGCTAGCTTTCCAGTTCTTTCATAA 2400
2401 TTCACAAATACCAAGTTTAACTAAGCAATTCGATTAACCTTTTCTTAGGTTAATAA 2460
2401 TTCACAAATACCAAGTTTAACTAAGCAATTCGATTAACCTTTTCTTAGGTTAATAA 2460
2461 AATAGAAGTATTTTCCAGGACACGGAGAAAAAGTTTCTAGGAAAGATACCTAGTGTG 2520
2461 AATAGAAGTATTTTCCAGGACACGGAGAAAAAGTTTCTAGGAAAGATACCTAGTGTG 2520
2521 TTGGTAGTCTATGAGAATAACATTTGTATATTAATACTAACTTTCTTTAGGTTGCTA 2580
2521 TTGGTAGTCTATGAGAATAACATTTGTATATTAATACTAACTTTCTTTAGGTTGCTA 2580
2581 TTCTCAATAATTTGCTATCCCAATATGATGTTATGTTCTTCAATATGAGTTAGTTTACTG 2640
2581 TTCTCAATAATTTGCTATCCCAATATGATGTTATGTTCTTCAATATGAGTTAGTTTACTG 2640
2641 ATGCTGTTATATGAAAAATACAGCAACCACTGATTTGTCGACATGCTCTACTGATATCTG 2700
2641 ATGCTGTTATATGAAAAATACAGCAACCACTGATTTGTCGACATGCTCTACTGATATCTG 2700
2701 GTAAGTGCCACAGATACATCTATATATTAATCAATAAATGAGTTAGTTTAAATTAAGT 2760
2701 GTAAGTGCCACAGATACATCTATATATTAATCAATAAATGAGTTAGTTTAAATTAAGT 2760
2761 TAGCAATTGATGCTTTCTCTATATTTCTTTGGCCAAAAGGCAAAAGTATTTCTCTTA 2820
2761 TAGCAATTGATGCTTTCTCTATATTTCTTTGGCCAAAAGGCAAAAGTATTTCTCTTA 2820
2821 AGTCTGGATTGCGGGTAATTTTGGGGCATGGGACCCATTTCTCATTTACGAGGCTCTG 2880
2821 AGTCTGGATTGCGGGTAATTTTGGGGCATGGGACCCATTTCTCATTTACGAGGCTCTG 2880
2881 GTGCCAGACAATAAGTAACTTATCTTAAATTTGGAGTTTACCAATTTGTAATAAGAG 2940
2881 GTGCCAGACAATAAGTAACTTATCTTAAATTTGGAGTTTACCAATTTGTAATAAGAG 2940
2941 TGACTAAACATATTTAATCAATTTGTAATAATTAATGTAATAATGCTATGTAATGTT 3000
2941 TGACTAAACATATTTAATCAATTTGTAATAATTAATGTAATAATGCTATGTAATGTT 3000
3001 GAGACTGTTATTTGGTAATTAAGAGTTGGTTTAAATTTGTAATTTCTCTCTTTTTCAG 3060
3001 GAGACTGTTATTTGGTAATTAAGAGTTGGTTTAAATTTGTAATTTCTCTCTTTTTCAG 3060
3061 CCCCCAAGCAATATGTAAGTATGATACA 3091
3061 CCCCCAAGCAATATGTAAGTATGATACA 3091

US-10-652-981-1
; Sequence 1, Application US/10652981
; Publication No. US2005074400A1
; GENERAL INFORMATION:
; APPLICANT: AGY Therapeutics
; APPLICANT: Meicher, Thorsten
; APPLICANT: Mueller, Sabine
; APPLICANT: Chin, Daniel
; TITLE OF INVENTION: USE OF PROTEIN BIOMOLECULAR TARGETS IN THE TREATMENT AND VISUALIZ
; FILE REFERENCE: 263/180 -- Preigleman -- AGY
; CURRENT APPLICATION NUMBER: US/10/652,981
; PRIOR FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/983,000A
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3091
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(3091)
; OTHER INFORMATION: PTP-zeta SMI exon 9 variant
; FEATURE:
; NAME/KEY: VARSPLIC
; LOCATION: (1262)..(1262)
; OTHER INFORMATION: Alternative splice site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1273)..(3091)
; OTHER INFORMATION: 3' Untranslated Region
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (148)..(1272)
; OTHER INFORMATION:
US-10-652-981-1

Query Match 100.0%; Score 3091; DB 21; Length 3091;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3091; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CACACATACGACGACGACGATCTCACTTCGATCTATACACTGGAGGATTTAAAAACAACAAA 60
Db 1 CACACATACGACGACGACGATCTCACTTCGATCTATACACTGGAGGATTTAAAAACAACAAA 60
Qy 61 CAAAAAAAACATTTCTTGGCTCCCTCTCCCTCTCTGAGAGCAGAGGAGCGCA 120
Db 61 CAAAAAAAACATTTCTTGGCTCCCTCTCCCTCTCTGAGAGCAGAGGAGCGCA 120
Qy 121 CGGCGAGGGGCGCAGACCGCTCTGGAATGGAATCTTAAAGCGTTTCTCGCTTGCATT 180
Db 121 CGGCGAGGGGCGCAGACCGCTCTGGAATGGAATCTTAAAGCGTTTCTCGCTTGCATT 180
Qy 181 CAGCTCTCTGTGTTTCCGCTGGATTTGGCTTAATGATACAGACAGACAGAGAAA 240
Db 181 CAGCTCTCTGTGTTTCCGCTGGATTTGGCTTAATGATACAGACAGACAGAGAAA 240
Qy 241 CTTGTTGAAGAGATTTGGCTGCTTATACAGGAGCACTGAATCAAAAAAATTTGGGGAAG 300
Db 241 CTTGTTGAAGAGATTTGGCTGCTTATACAGGAGCACTGAATCAAAAAAATTTGGGGAAG 300
Qy 301 AATATATCAACATGTAATAGCCCAAAACATCTCTATCAATATTGATGAAGATCTTACA 360
Db 301 AATATATCAACATGTAATAGCCCAAAACATCTCTATCAATATTGATGAAGATCTTACA 360
Qy 361 CAAGTAATGTGAATCTTAAAGAACTTTAAATTTTCAAGGTTTGGGATAAAACATCATTCGAA 420
Db 361 CAAGTAATGTGAATCTTAAAGAACTTTAAATTTTCAAGGTTTGGGATAAAACATCATTCGAA 420
Qy 421 AACACATTCATTCAATCACTGGGAAACAGTGGAAATTAATCTCACTAATGATACCGT 480
Db 421 AACACATTCATTCAATCACTGGGAAACAGTGGAAATTAATCTCACTAATGATACCGT 480

Db 421 AACACATTCATTCAACACACTGGGAAACACAGTGGAAATTAATCTCACTAATGACTACCGT 480
Qy 481 GTCAGCGGAGGAGTTTCAGAAATGGTGTAAAGCAAGACAGATAAATCTTTTCACTGGGGA 540
Db 481 GTCAGCGGAGGAGTTTCAGAAATGGTGTAAAGCAAGACAGATAAATCTTTTCACTGGGGA 540
Qy 541 AAATGCAATATGTCATCTGATGGATCAGAGCATAGTTTAGAAGCAAAATTTCCACTT 600
Db 541 AAATGCAATATGTCATCTGATGGATCAGAGCATAGTTTAGAAGCAAAATTTCCACTT 600
Qy 601 GAGATGCAAACTCTACTGCTGTGATGCGGACCGATTTTCAAGTTTGGAGAGCAGTCAAA 660
Db 601 GAGATGCAAACTCTACTGCTGTGATGCGGACCGATTTTCAAGTTTGGAGAGCAGTCAAA 660
Qy 661 GGAAGGGAAGTTAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAATTTTG 720
Db 661 GGAAGGGAAGTTAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAGAAATTTTG 720
Qy 721 GATTTCAAGCGATTAATGATGGAGTCGAAGTGTAGTCGTTTGGAGCAGCGCTGCT 780
Db 721 GATTTCAAGCGATTAATGATGGAGTCGAAGTGTAGTCGTTTGGAGCAGCGCTGCT 780
Qy 781 TTAGATCCATTCATACTCTGTAACCTTCTGCCAAACTCAACTGACAACTATTACATTTAC 840
Db 781 TTAGATCCATTCATACTCTGTAACCTTCTGCCAAACTCAACTGACAACTATTACATTTAC 840
Qy 841 AATGGCTCATTGACATCTCCTCCCTGACAGACACAGTTGACTGGATTTTAAAGAT 900
Db 841 AATGGCTCATTGACATCTCCTCCCTGACAGACACAGTTGACTGGATTTTAAAGAT 900
Qy 901 ACAGTTAGCATCTCGAAGCCAGTTGGCTGTTTGTGGAAGTCTTACATGCAACAA 960
Db 901 ACAGTTAGCATCTCGAAGCCAGTTGGCTGTTTGTGGAAGTCTTACATGCAACAA 960
Qy 961 TCTGGTTATGTCATGCTGATGACTACTTACAAACAAATTTTCGAGAGCAACAGTACAA 1020
Db 961 TCTGGTTATGTCATGCTGATGACTACTTACAAACAAATTTTCGAGAGCAACAGTACAA 1020
Qy 1021 TTCTCTAGACAGGTTTTCCTCATACACTGCGAAAGGAGAGATTCATGAAGCAGTTTGT 1080
Db 1021 TTCTCTAGACAGGTTTTCCTCATACACTGCGAAAGGAGAGATTCATGAAGCAGTTTGT 1080
Qy 1081 AGTTCAGAACCAAGAAATGTTCCAGCTGACCCAGAGATTTATCCAGCTTCTTGTGTACA 1140
Db 1081 AGTTCAGAACCAAGAAATGTTCCAGCTGACCCAGAGATTTATCCAGCTTCTTGTGTACA 1140
Qy 1141 TGGAAAGACCTCCAGTCCGTTTATGATACCATGATTGAGAGTTTGCAGTTTGTACCA 1200
Db 1141 TGGAAAGACCTCCAGTCCGTTTATGATACCATGATTGAGAGTTTGCAGTTTGTACCA 1200
Qy 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGAAATTTTGAAGAGATGGCTATCAAGACTTG 1260
Db 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGAAATTTTGAAGAGATGGCTATCAAGACTTG 1260
Qy 1261 GTTACTATATGATCAGTTGTTTACATAGGGTACATTTAATTTAATTTCCAGAGTAAG 1320
Db 1261 GTTACTATATGATCAGTTGTTTACATAGGGTACATTTAATTTAATTTCCAGAGTAAG 1320
Qy 1321 AACTTACAAATGGTTGTATATTTTCTCCATCTATTTTAGACTTTTATGGAAGGTGG 1380
Db 1321 AACTTACAAATGGTTGTATATTTTCTCCATCTATTTTAGACTTTTATGGAAGGTGG 1380
Qy 1381 GGTAGGCTGAGTATTTTAAATTTTAAATTTTAAATTTTAAATTTAGAGCTATCTAAATTA 1440
Db 1381 GGTAGGCTGAGTATTTTAAATTTTAAATTTTAAATTTTAAATTTAGAGCTATCTAAATTA 1440
Qy 1441 TGTTTTAAAGTTACATTTAATTTAAATGGATATCATAACTTTTCCACATTAACCTATAGA 1500
Db 1441 TGTTTTAAAGTTACATTTAATTTAAATGGATATCATAACTTTTCCACATTAACCTATAGA 1500
Qy 1501 GTAGATACATATGACTTATGAATCGGAGATCATTTAGTGTGGCCCTTCTTAAGATTTTCA 1560
Db 1501 GTAGATACATATGACTTATGAATCGGAGATCATTTAGTGTGGCCCTTCTTAAGATTTTCA 1560

Qy 1561 TTGTAGAAATAGTGCAGAAATCTCAGTGCCCTGATACATTTTATATTGTGTCTTCCATTAC 1620
Db 1561 TTGTAGAAATAGTGCAGAAATCTCAGTGCCCTGATACATTTTATATTGTGTCTTCCATTAC 1620
Qy 1621 GCTATATCAGCACAGGAAAGTAGAGTAGGGGACATACAAAGTCTCTTTTGTGACCCAAA 1680
Db 1621 GCTATATCAGCACAGGAAAGTAGAGTAGGGGACATACAAAGTCTCTTTTGTGACCCAAA 1680
Qy 1681 AAATTTTTCAGATAACAGCTGGGAAGTCATGATTTGGGTTCAGAACTTTTGGGAGTGAAGAAA 1740
Db 1681 AAATTTTTCAGATAACAGCTGGGAAGTCATGATTTGGGTTCAGAACTTTTGGGAGTGAAGAAA 1740
Qy 1741 ACATTTCTTACAAAAGATCCACCCCTGCTCCCTCCACAGCGCATGCGAATAAGTAC 1800
Db 1741 ACATTTCTTACAAAAGATCCACCCCTGCTCCCTCCACAGCGCATGCGAATAAGTAC 1800
Qy 1801 AGATTTCCCTTTTGTGGCTGAGCATGTGAGTATTAAACTTTTGTCTGTGGTAGGGAAGTGTG 1860
Db 1801 AGATTTCCCTTTTGTGGCTGAGCATGTGAGTATTAAACTTTTGTCTGTGGTAGGGAAGTGTG 1860
Qy 1861 GCCATAGATTAGGGTGTAGTTGACAAACCTTTCATCTGGATGTAGGTCAGAGAAAGTCCCCA 1920
Db 1861 GCCATAGATTAGGGTGTAGTTGACAAACCTTTCATCTGGATGTAGGTCAGAGAAAGTCCCCA 1920
Qy 1921 CTGCAGGTTAAAGGACACTGCACTCTGCACTCAGGCACTAGAGTCTCTGCAAGTCTCTGG 1980
Db 1921 CTGCAGGTTAAAGGACACTGCACTCTGCACTCAGGCACTAGAGTCTCTGCAAGTCTCTGG 1980
Qy 1981 AACCTGCATTTAAATAAAATGCACTATTAAATATTATTTTCATATCATGTGACAAATATG 2040
Db 1981 AACCTGCATTTAAATAAAATGCACTATTAAATATTATTTTCATATCATGTGACAAATATG 2040
Qy 2041 ATAAATTTTAGTAACCTTTTAAATTCAGTTTGCCTGGAAATGAGAGACAAATGACCTGG 2100
Db 2041 ATAAATTTTAGTAACCTTTTAAATTCAGTTTGCCTGGAAATGAGAGACAAATGACCTGG 2100
Qy 2101 AAAATCGTAATAATAGTAATAAAATGTTTATTTTCATAATTTACGTGAAGAGATAAT 2160
Db 2101 AAAATCGTAATAATAGTAATAAAATGTTTATTTTCATAATTTACGTGAAGAGATAAT 2160
Qy 2161 TCTATTACTGTTCTTGTGATATATATGTCAGGAAAGAGAGATAACTTAGTTGTTCACTTT 2220
Db 2161 TCTATTACTGTTCTTGTGATATATATGTCAGGAAAGAGAGATAACTTAGTTGTTCACTTT 2220
Qy 2221 TTCAATTTGCTCTGTTTGCATAATGCCCCCAATTTATTTGTCTAAATAATTAATTTTAA 2280
Db 2221 TTCAATTTGCTCTGTTTGCATAATGCCCCCAATTTATTTGTCTAAATAATTAATTTTAA 2280
Qy 2281 GTTTGTAGTACTAATTTTATGAAATTTGATGAGTTCTGGCTAAATGAACTTCCCTGAAC 2340
Db 2281 GTTTGTAGTACTAATTTTATGAAATTTGATGAGTTCTGGCTAAATGAACTTCCCTGAAC 2340
Qy 2341 TAAATCTGATTTTAAAGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
Db 2341 TAAATCTGATTTTAAAGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
Qy 2401 TTCAAAATACCAAGTTTAACTAAGCAAACTGATGATAAATTTTCTTAGGTTAATAA 2460
Db 2401 TTCAAAATACCAAGTTTAACTAAGCAAACTGATGATAAATTTTCTTAGGTTAATAA 2460
Qy 2461 AATAGAGTATTTTCCAGGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
Db 2461 AATAGAGTATTTTCCAGGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
Qy 2521 TTGTAGTCTCTATGAGATAAATTTTGTATAATTTTACTAAATTTTCTTAGGAGGAGGAGGAG 2580
Db 2521 TTGTAGTCTCTATGAGATAAATTTTGTATAATTTTACTAAATTTTCTTAGGAGGAGGAGGAG 2580
Qy 2581 TTCTCAATAAATTTTGTCTACCAATATAGTATGTTTCTTCAGATAGTAGCCATATGCACTA 2640
Db 2581 TTCTCAATAAATTTTGTCTACCAATATAGTATGTTTCTTCAGATAGTAGCCATATGCACTA 2640

QY 2641 ATGGCTTATATGGAATAACAGCAACCACTGATGTGACATGCTACTGATAATCCTG 2700
Db |||||
QY 2641 ATGGCTTATATGGAATAACAGCAACCACTGATGTGACATGCTACTGATAATCCTG 2700
Db |||||
QY 2701 GTAAGTGCCACAGATACATCTATATTAATCTCAATAATGAGGTAGTTAAATCTG 2760
Db |||||
QY 2701 GTAAGTGCCACAGATACATCTATATTAATCTCAATAATGAGGTAGTTAAATCTG 2760
Db |||||
QY 2761 TATGCAATGATGCTTCTCTCTATATCTTTTGGCCAAAGGCAAGTGAATTTCTCTTA 2820
Db |||||
QY 2761 TATGCAATGATGCTTCTCTCTATATCTTTTGGCCAAAGGCAAGTGAATTTCTCTTA 2820
Db |||||
QY 2821 AGTCTGGATGCGGGTAATTTTGGGCAATGGGACCACTTCTCAATCAGCAGGTCTG 2880
Db |||||
QY 2821 AGTCTGGATGCGGGTAATTTTGGGCAATGGGACCACTTCTCAATCAGCAGGTCTG 2880
Db |||||
QY 2881 GTGCCAGACAATGTAATCTTATCTTAATTTGGAGTTTACCAATTTGTAAATAAGAG 2940
Db |||||
QY 2881 GTGCCAGACAATGTAATCTTATCTTAATTTGGAGTTTACCAATTTGTAAATAAGAG 2940
Db |||||
QY 2941 TGAATAAACAATTTATTAATCTTGAATTAATCAATTAATGAAATTTGCTATGTAATGTT 3000
Db |||||
QY 2941 TGAATAAACAATTTATTAATCTTGAATTAATCAATTAATGAAATTTGCTATGTAATGTT 3000
Db |||||
QY 3001 GAGACTGTTATTTGGATAATTAAGAGTTGGTTAAATTTGTAATTTATTTCTCTTTTCAG 3060
Db |||||
QY 3001 GAGACTGTTATTTGGATAATTAAGAGTTGGTTAAATTTGTAATTTATTTCTCTTTTCAG 3060
Db |||||
QY 3061 CCCCCAAGCATATGTAAGTGTATATACA 3091
Db |||||
QY 3061 CCCCCAAGCATATGTAAGTGTATATACA 3091
Db |||||

RESULT 3

US-10-295-027-373
; Sequence 373, Application US/10295027
; Publication No. US2003023250A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Ebs Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 373
; LENGTH: 5367
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-373

Query Match 40.9%; Score 1263.2; DB 17; Length 5367;
Best Local Similarity 99.8%; Pred. No. 1.2e-257;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CACACATACGACGACGACGATCTCACTTCGATCTATACACTCGAGGATTTAAACAACAAA 60
Db |||||
QY 1 CACACATACGACGACGACGATCTCACTTCGATCTATACACTCGAGGATTTAAACAACAAA 60
Db |||||
QY 61 CAAAAAACAATTTCTTTCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db |||||
QY 61 CAAAAAACAATTTCTTTCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db |||||
QY 121 CGCGAGGGGCGCGAGACCGCTCTGGAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180
Db |||||
QY 121 CGCGAGGGGCGCGAGACCGCTCTGGAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180
Db |||||
QY 181 CAGCTCTCTCTGTTTGGCGCTCGATTTGGGCTAATGGATCTACAGACAACAGAGAAA 240
Db |||||
QY 181 CAGCTCTCTCTGTTTGGCGCTCGATTTGGGCTAATGGATCTACAGACAACAGAGAAA 240
Db |||||
QY 241 CTTGTTGAAGAGATTTGGCTCTCTATACAGGAGCAGTGAATCAAAAAAATTTGGGGAAG 300
Db |||||
QY 241 CTTGTTGAAGAGATTTGGCTCTCTATACAGGAGCAGTGAATCAAAAAAATTTGGGGAAG 300
Db |||||
QY 301 AAATATCCACATGTAATAGCCCAAAACAATCTCTATCAATATTTGATGAAGATCTTACA 360
Db |||||
QY 301 AAATATCCACATGTAATAGCCCAAAACAATCTCTATCAATATTTGATGAAGATCTTACA 360
Db |||||
QY 361 CAAGTAAATGTGAATCTTAAAGAACTTAAATTTCAAGGTTGGGATAAACATCATTTGAA 420
Db |||||
QY 361 CAAGTAAATGTGAATCTTAAAGAACTTAAATTTCAAGGTTGGGATAAACATCATTTGAA 420
Db |||||
QY 421 AACACATTCATTAACACACTGGGAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
Db |||||
QY 421 AACACATTCATTAACACACTGGGAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
Db |||||
QY 481 GTGAGCGGAGAGTTTCAAGAAATGCTTTTAAAGCAAGCAAGATACTTTTCACTGGGGA 540
Db |||||
QY 481 GTGAGCGGAGAGTTTCAAGAAATGCTTTTAAAGCAAGCAAGATACTTTTCACTGGGGA 540
Db |||||
QY 541 AAATGCAATATGTCATCTGATGGATCAGAGCATAGTTTAAAGGACAAAAAATTTTCCACT 600
Db |||||
QY 541 AAATGCAATATGTCATCTGATGGATCAGAGCATAGTTTAAAGGACAAAAAATTTTCCACT 600
Db |||||
QY 601 GAGATGCAAAATCTACTCTTTCGATGCGGACCGATTTTCAAGTTTGGAGGAGAGTCAAA 660
Db |||||
QY 601 GAGATGCAAAATCTACTCTTTCGATGCGGACCGATTTTCAAGTTTGGAGGAGAGTCAAA 660
Db |||||
QY 661 GGAAGAGGAGTTAAGAGCTTTTATCCATTTTGTGAGTTGGGACAGAGAAATTTT 720
Db |||||
QY 661 GGAAGAGGAGTTAAGAGCTTTTATCCATTTTGTGAGTTGGGACAGAGAAATTTT 720
Db |||||
QY 721 GATTTCAAGGAGTTAATGATGGAGTCAAGAGTGTAGTCTTTTGGAGGAGCGCTGCT 780
Db |||||
QY 721 GATTTCAAGGAGTTAATGATGGAGTCAAGAGTGTAGTCTTTTGGAGGAGCGCTGCT 780
Db |||||
QY 781 TTAGATCCATTCATCTGTTGAACCTTTCTGCCAACTCAACTGACAAGTATTACATTAC 840
Db |||||
QY 781 TTAGATCCATTCATCTGTTGAACCTTTCTGCCAACTCAACTGACAAGTATTACATTAC 840
Db |||||
QY 841 AATGGCTCATTTGACATCT 900
Db |||||
QY 841 AATGGCTCATTTGACATCT 900
Db |||||

```
QY 901 ACAGTTAGCATCTCGAAGCCAGTTGGCTGTTTTGTGAAGTTCTTACAAATCAACAA 960
Db 901 ACAGTTAGCATCTCGAAGCCAGTTGGCTGTTTTGTGAAGTTCTTACAAATCAACAA 960
QY 961 TCTGTTATGTGTCATGCTGATGGACTACTTACAAAACAATTTTCGAGAGCAACAGTACAA 1020
Db 961 TCTGTTATGTGTCATGCTGATGGACTACTTACAAAACAATTTTCGAGAGCAACAGTACAA 1020
QY 1021 TTCTCTAGACAGGTGTTTTCTCTATACACTGGAAGGAGAGATTCATGAAGCAGTTTGT 1080
Db 1021 TTCTCTAGACAGGTGTTTTCTCTATACACTGGAAGGAGAGATTCATGAAGCAGTTTGT 1080
QY 1081 AGTTCCAGAACCAAGAAATGTTTCAGGCTGACCCAGAGAAATATACAGGCTTCTTGTTACA 1140
Db 1081 AGTTCCAGAACCAAGAAATGTTTCAGGCTGACCCAGAGAAATATACAGGCTTCTTGTTACA 1140
QY 1141 TGGGAAGACCTCGAGTGGTTTATGATACCATGATTGAGAAGTTTGCAAGTTTGTGACCAG 1200
Db 1141 TGGGAAGACCTCGAGTGGTTTATGATACCATGATTGAGAAGTTTGCAAGTTTGTGACCAG 1200
QY 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGAAATTTTTCAGAGATGGCTATCAAGACTTG 1260
Db 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGAAATTTTTCAGAGATGGCTATCAAGACTTG 1260
QY 1261 GTAACAT 1268
Db 1261 GGTGCTAT 1268
```

RESULT 4

```
US-10-295-027-1044
; Sequence 1044, Application US/10295027
; Publication No. US2003023350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1044
```

```
; LENGTH: 5367
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-1044

Query Match      40.9%; Score 1263.2; DB 17; Length 5367;
Best Local Similarity 99.8%; Pred. No. 1.2e-257;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACACATACGACGACGATCTCATTTCGATCTATACACTGGAGGATTAACAAACAA 60
Db 1 CACACATACGACGACGATCTCATTTCGATCTATACACTGGAGGATTAACAAACAA 60
QY 61 CAAAAAACAATTTCTTCGTCCTCCCTCTCCACTCTGAGAACGAGGAGCCGCA 120
Db 61 CAAAAAACAATTTCTTCGTCCTCCCTCTCCACTCTGAGAACGAGGAGCCGCA 120
QY 121 CGCGAGGGGCGCGACGACGCTCTGGAATCGGAATCCTAAAGCGTTTCTCGCTTGCAT 180
Db 121 CGCGAGGGGCGCGACGACGCTCTGGAATCGGAATCCTAAAGCGTTTCTCGCTTGCAT 180
QY 181 CAGCTCCTCTGTGTTTTCGCGCTGGATTGGGCTAAATGGATACTACAGAACAGAGAAA 240
Db 181 CAGCTCCTCTGTGTTTTCGCGCTGGATTGGGCTAAATGGATACTACAGAACAGAGAAA 240
QY 241 CTGTTGAGAGATTGGCTGCTCTATACAGGAGCACTGAATCAAAAAAATTTGGGGAAG 300
Db 241 CTGTTGAGAGATTGGCTGCTCTATACAGGAGCACTGAATCAAAAAAATTTGGGGAAG 300
QY 301 AATATCCAACTGTAATAGCCCCAAAAACAATCTCTATCAATATTGATGAAGATCTTACA 360
Db 301 AATATCCAACTGTAATAGCCCCAAAAACAATCTCTATCAATATTGATGAAGATCTTACA 360
QY 361 CAAGTAAATGTGAATCTTAAAGAACTTTAAATTTCAAGGTTGGGATAAAACATCATTTG 420
Db 361 CAAGTAAATGTGAATCTTAAAGAACTTTAAATTTCAAGGTTGGGATAAAACATCATTTG 420
QY 421 AACACATTCATTCTAATACACTGGGAAAAACAGTGGAAATTAATCTCCTAATGACTACCG 480
Db 421 AACACATTCATTCTAATACACTGGGAAAAACAGTGGAAATTAATCTCCTAATGACTACCG 480
QY 481 GTACGCGAGGAGTTTTCAGAAATGTTTAAAGCAAGCAAGATAAATTTTCTCATCTGGGA 540
Db 481 GTACGCGAGGAGTTTTCAGAAATGTTTAAAGCAAGCAAGATAAATTTTCTCATCTGGGA 540
QY 541 AAATGCAATATGTCATCTGATGGATCAGAGCATAGTTTGAAGGACAAAAATTTCCACTT 600
Db 541 AAATGCAATATGTCATCTGATGGATCAGAGCATAGTTTGAAGGACAAAAATTTCCACTT 600
QY 601 GAGATGCAAAATCTACTCTTTCGCGACCGGATTTTCAAGTTTTCAGGAGCAAGTCAAA 660
Db 601 GAGATGCAAAATCTACTCTTTCGCGACCGGATTTTCAAGTTTTCAGGAGCAAGTCAAA 660
QY 661 GGAAGAGGGAAGTTAAGAGCTTTTATCCATTTTGTTCAGGTTGGGACAGAGAAATTTG 720
Db 661 GGAAGAGGGAAGTTAAGAGCTTTTATCCATTTTGTTCAGGTTGGGACAGAGAAATTTG 720
QY 721 GATTTCAAAAGCGATTATTGATGGAGTCCAAAGTGTGTAGTCTGTTTGGAGAGCAGCTGCT 780
Db 721 GATTTCAAAAGCGATTATTGATGGAGTCCAAAGTGTGTAGTCTGTTTGGAGAGCAGCTGCT 780
QY 781 TTAGATCCATTCTACTCTTTCGCGACCGGATTTTCAAGTTTTCAGGAGCAAGTCAAA 840
Db 781 TTAGATCCATTCTACTCTTTCGCGACCGGATTTTCAAGTTTTCAGGAGCAAGTCAAA 840
QY 841 AATGGCTCATTGACATCTCTCCCTGCACAGACACAGTTGACTGGATTTGTTTAAAGAT 900
Db 841 AATGGCTCATTGACATCTCTCCCTGCACAGACACAGTTGACTGGATTTGTTTAAAGAT 900
QY 901 ACAGTTAGCATCTCTGAAGCCAGTTGGCTGTTTTGTGAAGTTCTTACAAATCAACAA 960
Db 901 ACAGTTAGCATCTCTGAAGCCAGTTGGCTGTTTTGTGAAGTTCTTACAAATCAACAA 960
```



```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 381
; LENGTH: 7941
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-381

Query Match      40.9%; Score 1263.2; DB 17; Length 7941;
Best Local Similarity 99.8%; Pred. No. 1.4e-257;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACACATACGACGACGACGATCTCCTCGATCTATACACTGAGGATTTAAACAAACAAA 60
DB 1 CACACATACGACGACGACGATCTCCTCGATCTATACACTGAGGATTTAAACAAACAAA 60
QY 61 CAAAAAACAATTCCTTCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 120
DB 61 CAAAAAACAATTCCTTCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 120
QY 121 CGGAGAGGGGGCGGACCGCTCTGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 180
DB 121 CGGAGAGGGGGCGGACCGCTCTGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 180
QY 181 CAGCTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 181 CAGCTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 CTTGTTGAAGAGATGGCTGCTCTCTATACAGGAGCAGTGAATCAAAAAAATTTGGGAAAG 300
DB 241 CTTGTTGAAGAGATGGCTGCTCTCTATACAGGAGCAGTGAATCAAAAAAATTTGGGAAAG 300
QY 301 AAATATCCAAATGATAGCCCAAAACAAATCTCTATCAATATGATGAAGATCTTACA 360
DB 301 AAATATCCAAATGATAGCCCAAAACAAATCTCTATCAATATGATGAAGATCTTACA 360
QY 361 CAAGTAAATGTGAATCTTAAAGAACTTAAATTCAGGGTTCGGATATAAATCAATCTGGA 420
DB 361 CAAGTAAATGTGAATCTTAAAGAACTTAAATTCAGGGTTCGGATATAAATCAATCTGGA 420
QY 421 AACACATTCATTAACAACATGGGAAACACAGTGGAAATTAATCTCAATGATGATACCGT 480
DB 421 AACACATTCATTAACAACATGGGAAACACAGTGGAAATTAATCTCAATGATGATACCGT 480
QY 481 CTCAGCGAGAGGATTCAGAAATGCTGTTTAAAGCAAGCAGATAACTTTTCACTGGGGA 540
DB 481 CTCAGCGAGAGGATTCAGAAATGCTGTTTAAAGCAAGCAGATAACTTTTCACTGGGGA 540
QY 541 AAATGCAATATGTCATCTGATGGATCAGAGCATAGTTTGAAGGACAAAAATTTCCACTT 600
DB 541 AAATGCAATATGTCATCTGATGGATCAGAGCATAGTTTGAAGGACAAAAATTTCCACTT 600
QY 601 GAGATGCAAAATCTACTGCTTTGATGCGGACCGATTTTCAAGTTTGAAGGAGCAGTCAAA 660
DB 601 GAGATGCAAAATCTACTGCTTTGATGCGGACCGATTTTCAAGTTTGAAGGAGCAGTCAAA 660
QY 661 GGAAAGGGAATTAAGAGCTTTATCCATTTTGTGAGGTTGGGACGAGGAAATTTG 720
DB 661 GGAAAGGGAATTAAGAGCTTTATCCATTTTGTGAGGTTGGGACGAGGAAATTTG 720
QY 721 GATTTCAAAGCGATTTATGATGGAGTCCAAAGTGTAGTCTGTTTGGGAGCAGGCTGCT 780
DB 721 GATTTCAAAGCGATTTATGATGGAGTCCAAAGTGTAGTCTGTTTGGGAGCAGGCTGCT 780
QY 781 TTAGATCCATTCATCTGTTGAACCTTCTGCAAACTCAACTGACGATTTATCAATTTAC 840
DB 781 TTAGATCCATTCATCTGTTGAACCTTCTGCAAACTCAACTGACGATTTATCAATTTAC 840
QY 841 AATGGCTCATTCAGATCTCCTCCCTGACACACAGTTGATGCTGATGTTTAAAGAT 900
DB 841 AATGGCTCATTCAGATCTCCTCCCTGACACACAGTTGATGCTGATGTTTAAAGAT 900
QY 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGCTGTTTGTGAAGTCTTACATGCAACAA 960
```

```
DB 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGCTGTTTGTGAAGTCTTACAAATGCAACAA 960
QY 961 TCTGGTTATGTCATGCTGATGGACTACTTACAAAAAATTTTTCAGAGCAACAGTACAAAG 1020
DB 961 TCTGGTTATGTCATGCTGATGGACTACTTACAAAAAATTTTTCAGAGCAACAGTACAAAG 1020
QY 1021 TTCTCTAGACAGGTGTTTCTCTCATACACCTGGAAGCAAGAGATTTCAATGAAGCAGTTTGT 1080
DB 1021 TTCTCTAGACAGGTGTTTCTCTCATACACCTGGAAGCAAGAGATTTCAATGAAGCAGTTTGT 1080
QY 1081 AGTTCAAGAACACAGAAAATGTTTCAGGCTGACCCAGAGAAATATACAGCCCTTCTTGTTACA 1140
DB 1081 AGTTCAAGAACACAGAAAATGTTTCAGGCTGACCCAGAGAAATATACAGCCCTTCTTGTTACA 1140
QY 1141 TGGAAAGACCTCGAGTCGTTTATGATACCATGATGAGAAGTTTGAGTTTGTGATACAG 1200
DB 1141 TGGAAAGACCTCGAGTCGTTTATGATACCATGATGAGAAGTTTGAGTTTGTGATACAG 1200
QY 1201 CAGTTGATGGAGGAGGACCAACCAAGCATGAAATTTTGGACAGATGGCTATCAAGACTTG 1260
DB 1201 CAGTTGATGGAGGAGGACCAACCAAGCATGAAATTTTGGACAGATGGCTATCAAGACTTG 1260
QY 1261 GTAACTAT 1268
DB 1261 GGTGCTAT 1268

RESULT 9
US-10-843-641A-4839
; Sequence 4839, Application US/10843641A
; Publication No. US2005006454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4839
; LENGTH: 7941
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-4839

Query Match      40.9%; Score 1263.2; DB 21; Length 7941;
Best Local Similarity 99.8%; Pred. No. 1.4e-257;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACACATACGACGACGACGATCTCCTCGATCTATACACTGAGGATTTAAACAAACAAA 60
```


QY 361 CAAGTAAATGTGAATCTTAAGAACTTAAATTTTCAGGGTTGGGATATAACATCATTTGGAA 420
Db 361 CAAGTAAATGTGAATCTTAAGAACTTAAATTTTCAGGGTTGGGATATAACATCATTTGGAA 420
QY 421 AACACATTCATTCAATACACATCGGGAACACAGTGGAAATTAATCTCACTAATGACTACCGT 480
Db 421 AACACATTCATTCAATACACATCGGGAACACAGTGGAAATTAATCTCACTAATGACTACCGT 480
QY 481 GTCAGCGAGGAGTTTCAGAAATGGTGTTTAAAGCAACAGATAAATCTTTTCACCTGGGGA 540
Db 481 GTCAGCGAGGAGTTTCAGAAATGGTGTTTAAAGCAACAGATAAATCTTTTCACCTGGGGA 540
QY 541 AAATGCATATGTCTATCTGATGGATCAGAGCATAGTTTAGAGACAAATAATTTCCACTT 600
Db 541 AAATGCATATGTCTATCTGATGGATCAGAGCATAGTTTAGAGACAAATAATTTCCACTT 600
QY 601 GAGATGCAAACTACTGCTTTGATGCGGACCGATTTTCAAGTTTGTAGGAAGCAGTCAAA 660
Db 601 GAGATGCAAACTACTGCTTTGATGCGGACCGATTTTCAAGTTTGTAGGAAGCAGTCAAA 660
QY 661 GGAAGGGAAGTTAAGAGCTTTATCCATTTTGTGAGTTGGGACAGAAATAATTTG 720
Db 661 GGAAGGGAAGTTAAGAGCTTTATCCATTTTGTGAGTTGGGACAGAAATAATTTG 720
QY 721 GATTTCAAGCGATTAATGTGAGTGGAGTGAAGTGTAGTCTGTTTGGGAAGCAGCTGCT 780
Db 721 GATTTCAAGCGATTAATGTGAGTGGAGTGAAGTGTAGTCTGTTTGGGAAGCAGCTGCT 780
QY 781 TTAGATCCATTCATCACTGTTGAACCTTCTGCGCAACTCAACTGACAACTATTATACATTAC 840
Db 781 TTAGATCCATTCATCACTGTTGAACCTTCTGCGCAACTCAACTGACAACTATTATACATTAC 840
QY 841 AATGCTCATTGACATCTCTCTCCCTGACAGACAGATTTGATGATTTTAAAGAT 900
Db 841 AATGCTCATTGACATCTCTCTCCCTGACAGACAGATTTGATGATTTTAAAGAT 900
QY 901 ACAGTTAGCATCTCTGAAAGCGAGTTGGCTGTTTTTGTGAAGTCTTTACAAATGCACAA 960
Db 901 ACAGTTAGCATCTCTGAAAGCGAGTTGGCTGTTTTTGTGAAGTCTTTACAAATGCACAA 960
QY 961 TCTGTTATGTATGCTGATGACTACTTACAAACAATTTTCAGAGACAAACAGTACAA 1020
Db 961 TCTGTTATGTATGCTGATGACTACTTACAAACAATTTTCAGAGACAAACAGTACAA 1020
QY 1021 TTCTCTAGACAGTGTCTTCTCTATACACTGGAAGGAGAGATTCATGAAGCAGTTTGT 1080
Db 1021 TTCTCTAGACAGTGTCTTCTCTATACACTGGAAGGAGAGATTCATGAAGCAGTTTGT 1080
QY 1081 AGTTAGAACAGAAATGTTTCAGGCTGACCCAGAGAAATATACAGGCTTCTGTTTACA 1140
Db 1081 AGTTAGAACAGAAATGTTTCAGGCTGACCCAGAGAAATATACAGGCTTCTGTTTACA 1140
QY 1141 TGGAAAGACCTTCGAGTGGTTTATGATACCATGATTTGAGAAAGTTTGAGCTTTTGTAC 1200
Db 1141 TGGAAAGACCTTCGAGTGGTTTATGATACCATGATTTGAGAAAGTTTGAGCTTTTGTAC 1200
QY 1201 CAGTTGGATGGAGAGGACAAACCAAGCATGAATTTTGAAGATGGCTATCAAGACTTG 1260
Db 1201 CAGTTGGATGGAGAGGACAAACCAAGCATGAATTTTGAAGATGGCTATCAAGACTTG 1260
QY 1261 GTAACAT 1268
Db 1261 GGTGCTAT 1268

RESULT 11
US-10-936-626-32
; Sequence 32, Application US/10936626
; Publication No. US2005010664A1
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan

; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepfen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor
; FILE REFERENCE: P5001R1P1
; CURRENT APPLICATION NUMBER: US/10/936,626
; CURRENT FILING DATE: 2004-09-08
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 32
; LENGTH: 7941
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-936-626-32

Query Match 40.9%; Score 1263.2; DB 21; Length 7941;
Best Local Similarity 99.8%; Pred. No. 1.4e-257;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACACATACGACCGCAGATCTCACTTCGATCTATACACTGGAGATTAACACAAACAAA 60
Db 1 CACACATACGACCGCAGATCTCACTTCGATCTATACACTGGAGATTAACACAAACAAA 60
QY 61 CAAAAAACAATTTCTTCGCTCCCTCCCTCTCCACTCTGAGAACGAGAGCGCGCA 120
Db 61 CAAAAAACAATTTCTTCGCTCCCTCCCTCTCCACTCTGAGAACGAGAGCGCGCA 120
QY 121 CGGCGAGGCGCGCAGACCGCTCTGGAATCGGAATCTTAAAGCGTTTCTCGCTTGCATT 180
Db 121 CGGCGAGGCGCGCAGACCGCTCTGGAATCGGAATCTTAAAGCGTTTCTCGCTTGCATT 180
QY 181 CAGCTCTCTGTGTGTTTCCGCTGGATTTGGCTTAATGATACACACACAGAGAAA 240
Db 181 CAGCTCTCTGTGTGTTTCCGCTGGATTTGGCTTAATGATACACACACAGAGAAA 240
QY 241 CTTGTTGAAGAGATTTGGCTGCTCTATACAGGAGCACTGAATCAAAAAAATTTGGGGAAG 300
Db 241 CTTGTTGAAGAGATTTGGCTGCTCTATACAGGAGCACTGAATCAAAAAAATTTGGGGAAG 300
QY 301 AAATATCCACATGTATAGCCCAAAACAATCTCTATCAATATGATGAAGATCTTACA 360
Db 301 AAATATCCACATGTATAGCCCAAAACAATCTCTATCAATATGATGAAGATCTTACA 360
QY 361 CAAGTAAATGTGAATCTTAAGAACTTAAATTTTCAGGGTTGGGATATAACATCATTTGAA 420
Db 361 CAAGTAAATGTGAATCTTAAGAACTTAAATTTTCAGGGTTGGGATATAACATCATTTGAA 420
QY 421 AACACATTCATTCAATACACATCGGGAACACAGTGGAAATTAATCTCACTAATGACTACCGT 480

```
Db 421 AACACATTCATTCATAACACCTGGGAAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
Qy 481 GTCAGCGGAGGAGTTTCAGAAATGTTTAAAGCAAGCAAGATAACTTTTCACTGGGGA 540
Db 481 GTCAGCGGAGGAGTTTCAGAAATGTTTAAAGCAAGCAAGATAACTTTTCACTGGGGA 540
Qy 541 AAATCAATATGTCATCTGATGGATCAGAGCATAGTTTAGAGGACAAAATTTTCCATT 600
Db 541 AAATCAATATGTCATCTGATGGATCAGAGCATAGTTTAGAGGACAAAATTTTCCATT 600
Qy 601 GAGATGCAATCTACTGCTTTGATCGGACCGAATTTCAAGTTTGGAGAGCAGTCAAA 660
Db 601 GAGATGCAATCTACTGCTTTGATCGGACCGAATTTCAAGTTTGGAGAGCAGTCAAA 660
Qy 661 GGAAGGAGGAGTTAAGAGCTTTATCCATTTTGTGAGTTGGGACAGAGAAATTTG 720
Db 661 GGAAGGAGGAGTTAAGAGCTTTATCCATTTTGTGAGTTGGGACAGAGAAATTTG 720
Qy 721 GATTTCAAGCGGATTAATGATGAGTCGAAAGTGTAGTCTGTTTGGGAGCAGCTGCT 780
Db 721 GATTTCAAGCGGATTAATGATGAGTCGAAAGTGTAGTCTGTTTGGGAGCAGCTGCT 780
Qy 781 TTGATTCATTCATCTGTTGAACTCTGCGCAAACTCAACTGACAAAGTATTATCAATTAC 840
Db 781 TTGATTCATTCATCTGTTGAACTCTGCGCAAACTCAACTGACAAAGTATTATCAATTAC 840
Qy 841 AATGCTCATTTGACATCTCCCTCGCACACAGTTGACTGATTTGTTTAAAGAT 900
Db 841 AATGCTCATTTGACATCTCCCTCGCACACAGTTGACTGATTTGTTTAAAGAT 900
Qy 901 ACAGTTAGCATCTCGAAAGCCAGTTGGCTGTTTGTGAAAGTTCTTACAAATGCAACAA 960
Db 901 ACAGTTAGCATCTCGAAAGCCAGTTGGCTGTTTGTGAAAGTTCTTACAAATGCAACAA 960
Qy 961 TCTGTTATGTCATGCTGATGAGTACTTACAAACAAATTTTCGAGAGCAACAGTACAA 1020
Db 961 TCTGTTATGTCATGCTGATGAGTACTTACAAACAAATTTTCGAGAGCAACAGTACAA 1020
Qy 1021 TTCTCTACAGAGTGTTCCTCATACTGAGAGGAGAGATTCATGAGAGCAGTTTGT 1080
Db 1021 TTCTCTACAGAGTGTTCCTCATACTGAGAGGAGAGATTCATGAGAGCAGTTTGT 1080
Qy 1081 AGTTCAGAACCCAGAAAATGTTTCAGGCTGACCCAGAGAAATATACAGAGCTTCTGTTTACA 1140
Db 1081 AGTTCAGAACCCAGAAAATGTTTCAGGCTGACCCAGAGAAATATACAGAGCTTCTGTTTACA 1140
Qy 1141 TGAGAAAGACCTCGAGTCTGTTATGATACCATGATGAGAGTTTGGAGTTTGTACCA 1200
Db 1141 TGAGAAAGACCTCGAGTCTGTTATGATACCATGATGAGAGTTTGGAGTTTGTACCA 1200
Qy 1201 CAGTTGGATGAGAGGACCAACCAAGCATGAAATTTTTCAGAGTGGCTATCAAGACTTG 1260
Db 1201 CAGTTGGATGAGAGGACCAACCAAGCATGAAATTTTTCAGAGTGGCTATCAAGACTTG 1260
Qy 1261 GTAACAT 1268
Db 1261 GGTGCTAT 1268
```

RESULT 12

```
US-10-938-061-32
; Sequence 32, Application US/10938061
; Publication No. US20050107595A1
; GENERAL INFORMATION:
; APPLICANT: Cairns, Belinda
; APPLICANT: Chen, Ruihuan
; APPLICANT: Prantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Koepen, Hartmut
; APPLICANT: Phillips, Heidi S.
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan D.
```

```
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; APPLICANT: Sakanaka, Chie
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Reed Chae J.
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; FILE REFERENCE: P5001R1P1B
; CURRENT APPLICATION NUMBER: US/10/938,061
; CURRENT FILING DATE: 2004-09-10
; PRIOR APPLICATION NUMBER: US 10/872,991
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/872,972
; PRIOR FILING DATE: 2004-06-21
; PRIOR APPLICATION NUMBER: US 10/241,220
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/299,500
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 60/301,880
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/323,268
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/557,116
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US 60/598,899
; PRIOR FILING DATE: 2004-08-04
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 32
; LENGTH: 7941
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-938-061-32
```

```
Query Match 40.9%; Score 1263.2; DB 21; Length 7941;
Best Local Similarity 99.8%; Pred. No. 1.4e-257;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CACACATACGACGACGATCTCACTTCGATCTATACACTCGAGAGATTAACAAACAAA 60
Db 1 CACACATACGACGACGATCTCACTTCGATCTATACACTCGAGAGATTAACAAACAAA 60
Qy 61 CAAAAAAAACATTTCTTCCCTCCCTCCCTCTCCACTCTGAGAGCAGGAGCGCA 120
Db 61 CAAAAAAAACATTTCTTCCCTCCCTCCCTCTCCACTCTGAGAGCAGGAGCGCA 120
Qy 121 CGGCGGGGCGCGCAGACCGTCTGGAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180
Db 121 CGGCGGGGCGCGCAGACCGTCTGGAATGCGAATCTTAAAGCGTTTCTCGCTTGCATT 180
Qy 181 CAGCTCCTCTGTGTTTCCGCTGAGTTGGGCTAATGAGTACTACAGACAAACAGAGAAA 240
Db 181 CAGCTCCTCTGTGTTTCCGCTGAGTTGGGCTAATGAGTACTACAGACAAACAGAGAAA 240
Qy 241 CTGTTGAGAGATTTGGCTGCTTATACAGAGCAGTGAATCAAAAAATTTGGGGAAG 300
Db 241 CTGTTGAGAGATTTGGCTGCTTATACAGAGCAGTGAATCAAAAAATTTGGGGAAG 300
Qy 301 AATATATCAACATGTAATAGCCAAAAACAATCTCTATCAATATTGATGAAGATCTTACA 360
Db 301 AATATATCAACATGTAATAGCCAAAAACAATCTCTATCAATATTGATGAAGATCTTACA 360
Qy 361 CAAGTAATGTGAATCTTAAAGAACTTTAAATTTTTCAGGTTGGGATAAAAAATCAATGGA 420
Db 361 CAAGTAATGTGAATCTTAAAGAACTTTAAATTTTTCAGGTTGGGATAAAAAATCAATGGA 420
Qy 421 AACACATTCATTCATAACACTGGGAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
Db 421 AACACATTCATTCATAACACTGGGAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
```

```
Qy 481 CTCAGCGGAGGAGTTTCAGAAATGGTGTAAAGCAACAGATATACTTTTCACTGGGA 540
Db 481 CTCAGCGGAGGAGTTTCAGAAATGGTGTAAAGCAACAGATATACTTTTCACTGGGA 540
Qy 541 AAATGCAATATGTCATCTGATGGATCAGAGCATAGTTTAGAAGACAAAAATTTCCACTT 600
Db 541 AAATGCAATATGTCATCTGATGGATCAGAGCATAGTTTAGAAGACAAAAATTTCCACTT 600
Qy 601 GAGATGCAAACTACTGCTGTATGTCGGACCGATTTTCAAGTTTGGAGAGCAGTCAA 660
Db 601 GAGATGCAAACTACTGCTGTATGTCGGACCGATTTTCAAGTTTGGAGAGCAGTCAA 660
Qy 661 GGAAGAGGAGTAAAGAGCTTTTCCATTTTGTGAGGTTGGACAGAGAAATTTG 720
Db 661 GGAAGAGGAGTAAAGAGCTTTTCCATTTTGTGAGGTTGGACAGAGAAATTTG 720
Qy 721 GATTTCAAAGCGATTTATGATGGAGTCGAAAGTGTAGTCCGTTTGGGAAGCAGCGTCT 780
Db 721 GATTTCAAAGCGATTTATGATGGAGTCGAAAGTGTAGTCCGTTTGGGAAGCAGCGTCT 780
Qy 781 TTAGATCAATTCATCTGTTGAACTTCTGCGAAACTCAACTGACAGTATATCAATTAC 840
Db 781 TTAGATCAATTCATCTGTTGAACTTCTGCGAAACTCAACTGACAGTATATCAATTAC 840
Qy 841 AATGCTCATTTGACATCTCCTCCCTGCACACACAGTTGACTGGATTTTAAAGAT 900
Db 841 AATGCTCATTTGACATCTCCTCCCTGCACACACAGTTGACTGGATTTTAAAGAT 900
Qy 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTGTGAAGTTCTTACAATGCAACAA 960
Db 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTTTGTGAAGTTCTTACAATGCAACAA 960
Qy 961 TCTGTTATGTCATGCTGATGGAGTACTTACAAAACAAATTTTCGAGAGCAACAGTACAAG 1020
Db 961 TCTGTTATGTCATGCTGATGGAGTACTTACAAAACAAATTTTCGAGAGCAACAGTACAAG 1020
Qy 1021 TTCTCTACAGAGTGTGTTTCTCTATACACTGGAAGGAGAGATTTCATGAGCAGTTGT 1080
Db 1021 TTCTCTACAGAGTGTGTTTCTCTATACACTGGAAGGAGAGATTTCATGAGCAGTTGT 1080
Qy 1081 AGTTTCAAGAACAGAAATGTTTCAGGCTGACCCAGAGAAATATACAGCCTTTCTTGTACA 1140
Db 1081 AGTTTCAAGAACAGAAATGTTTCAGGCTGACCCAGAGAAATATACAGCCTTTCTTGTACA 1140
Qy 1141 TGGGAAAGACCTCGAGTGGTTATGATACCATGATGAGAGTTTGCAGTTTGTACAG 1200
Db 1141 TGGGAAAGACCTCGAGTGGTTATGATACCATGATGAGAGTTTGCAGTTTGTACAG 1200
Qy 1201 CAGTTGGATGGAGGAGGACCAACCAAGCATCAATTTTTCAGAGATGGCTATCAAGACTTG 1260
Db 1201 CAGTTGGATGGAGGAGGACCAACCAAGCATCAATTTTTCAGAGATGGCTATCAAGACTTG 1260
Qy 1261 GTAACTAT 1268
Db 1261 GGTGCTAT 1268
```

RESULT 13

```
US-10-295-027-371
; Sequence 371, Application US/10295027
; Publication No. US2003023250A1
; GENERAL INFORMATION:
; APPLICANT: Afat, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gieh, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hetezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Susan R.
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
```

```
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-0125000US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 371
; LENGTH: 7944
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-371
```

```
Query Match 40.9%; Score 1263.2; DB 17; Length 7944;
Best Local Similarity 99.8%; Pred. No. 1.4e-257;
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 CACACATACCGACGACGATCTCACTTCGATCTATACACTGGAGGATTAACACACAAA 60
Db 1 CACACATACCGACGACGATCTCACTTCGATCTATACACTGGAGGATTAACACACAAA 60
Qy 61 CAAAAAACAATTTCTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 120
Db 61 CAAAAAACAATTTCTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 120
Qy 121 CGGCGAGGCGCGACAGCCCTCTGGAATCGGAATCCTAAAGCGTTTCTCGCTTGCATT 180
Db 121 CGGCGAGGCGCGACAGCCCTCTGGAATCGGAATCCTAAAGCGTTTCTCGCTTGCATT 180
Qy 181 CAGCTCCTCTGTGTTGCGCGCTGGATTGGGCTTAATGGATCTACAGACACAGAGAAA 240
Db 181 CAGCTCCTCTGTGTTGCGCGCTGGATTGGGCTTAATGGATCTACAGACACAGAGAAA 240
Qy 241 CTGTTTGAAGAGATTGGCTGCTCTATACAGGAGCACTGAATCAAAAAAATTTGGGHAAG 300
Db 241 CTGTTTGAAGAGATTGGCTGCTCTATACAGGAGCACTGAATCAAAAAAATTTGGGHAAG 300
Qy 301 AAATATCCAACTGTAATAGCCCAAAACAATCTCCTATCAATATTTGATGAAGATCTTACA 360
Db 301 AAATATCCAACTGTAATAGCCCAAAACAATCTCCTATCAATATTTGATGAAGATCTTACA 360
Qy 361 CAAGTAATGTGAATCTTAAGAACTTAATTTTCAAGGTTGGGATTAACATCAATTGGAA 420
Db 361 CAAGTAATGTGAATCTTAAGAACTTAATTTTCAAGGTTGGGATTAACATCAATTGGAA 420
Qy 421 AACACATTCTATCACTACACTGGGAAAAAGTGGAAATTAATCTCACTAATGACTACCGT 480
Db 421 AACACATTCTATCACTACACTGGGAAAAAGTGGAAATTAATCTCACTAATGACTACCGT 480
Qy 481 GTACGCGGAGGAGTTTCAGAAATGGTGTAAAGCAAGAGATAAATTTTCTACTGGGGA 540
Db 481 GTACGCGGAGGAGTTTCAGAAATGGTGTAAAGCAAGAGATAAATTTTCTACTGGGGA 540
```



```
Qy 601 GAGATGCAAACTACTGCTTTGATGCGGACCGATTTTCAAGTTTGTAGGAAGCAGTCAAA 660
Db |||||
Qy 601 GAGATGCAAACTACTGCTTTGATGCGGACCGATTTTCAAGTTTGTAGGAAGCAGTCAAA 660
Db |||||
Qy 661 GGAAAGGGAAGTTAAAGAGCTTTTATCCATTTTGTGTAGAGTTGGGACAGAGAAAATTTG 720
Db |||||
Qy 661 GGAAAGGGAAGTTAAAGAGCTTTTATCCATTTTGTGTAGAGTTGGGACAGAGAAAATTTG 720
Db |||||
Qy 721 GATTTCAAAGCGATTATTTGATGGAGTGAAGAGTTAGTCTGTTTGGGAAGCAGCTGCT 780
Db |||||
Qy 721 GATTTCAAAGCGATTATTTGATGGAGTGAAGAGTTAGTCTGTTTGGGAAGCAGCTGCT 780
Db |||||
Qy 781 TTAGATCCATTATCATCTCTGCAACCTCTGCAAACTCAACTCAACTCAACTCAACTCAACT 840
Db |||||
Qy 781 TTAGATCCATTATCATCTCTGCAACCTCTGCAAACTCAACTCAACTCAACTCAACTCAACT 840
Db |||||
Qy 841 AATGGCTCATTTGACATCTCTCCCTGCAACAGTGTGACTGGATTTGTTTAAAGAT 900
Db |||||
Qy 841 AATGGCTCATTTGACATCTCTCCCTGCAACAGTGTGACTGGATTTGTTTAAAGAT 900
Db |||||
Qy 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTGTGGAAGTTCTTAAAGAT 960
Db |||||
Qy 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTGTGGAAGTTCTTAAAGAT 960
Db |||||
Qy 961 TCTGTTATGTCATGCTGATGAGTCTTATCAAAACAAATTTTCGAGAGCAACAGTACAA 1020
Db |||||
Qy 961 TCTGTTATGTCATGCTGATGAGTCTTATCAAAACAAATTTTCGAGAGCAACAGTACAA 1020
Db |||||
Qy 1021 TTCTCTAGACAGGTGTTTCTCTCATACACTGCGAAAGGAGAGATTCATGAAGCAGTTTGT 1080
Db |||||
Qy 1021 TTCTCTAGACAGGTGTTTCTCTCATACACTGCGAAAGGAGAGATTCATGAAGCAGTTTGT 1080
Db |||||
Qy 1081 AGTTCAGAACAGAAAATGTTTCAGCTGACCCAGAGAAATTAACAGCCTTCTTGTACAA 1140
Db |||||
Qy 1081 AGTTTCAAGAACAGAAAATGTTTCAGCTGACCCAGAGAAATTAACAGCCTTCTTGTACAA 1140
Db |||||
Qy 1141 TGGGAAGACCTCGAGTCTGTTTATGATACCATGATTGAGAGTTTGCAGTTTGTACCA 1200
Db |||||
Qy 1141 TGGGAAGACCTCGAGTCTGTTTATGATACCATGATTGAGAGTTTGCAGTTTGTACCA 1200
Db |||||
Qy 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGAATTTTTCAGAGATGGCTATCAAGACTTG 1260
Db |||||
Qy 1261 GTAACTAT 1268
Db |||||
Qy 1261 GGTGCTAT 1268
Db |||||
```

RESULT 15

```
US-10-852-335A-6
; Sequence 6, Application US/10852335A
; Publication No. US20050112129A1
; GENERAL INFORMATION:
; APPLICANT: HEIDI S. PHILLIPS
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; FILE OF INVENTION: Treatment of Tumors of Glial Origin
; FILE REFERENCE: P5103R1-US
; CURRENT APPLICATION NUMBER: US/10/852,335A
; CURRENT FILING DATE: 2004-05-24
; PRIOR APPLICATION NUMBER: US 60/548,299
; PRIOR FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/473,238
; PRIOR FILING DATE: 2003-05-23
; NUMBER OF SEQ ID NOS: 190
; SEQ ID NO 6
; LENGTH: 7944
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-852-335A-6
```

Query Match 40.9%; Score 1263.2; DB 21; Length 7944;
Best Local Similarity 99.8%; Pred. No. 1.4e-257;

```
Matches 1265; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CACACATACGACGACGACGATCTCACTTCGATCTATATACACTGGAGGATTAACAAACAAA 60
Db |||||
Qy 1 CACACATACGACGACGACGATCTCACTTCGATCTATATACACTGGAGGATTAACAAACAAA 60
Db |||||
Qy 61 CAAAAAACAATTTCTTCCTCCCTCTCTCCACTCTCTGAGAAGCAGAGGAGCCGCA 120
Db |||||
Qy 61 CAAAAAACAATTTCTTCCTCCCTCTCTCCACTCTCTGAGAAGCAGAGGAGCCGCA 120
Db |||||
Qy 121 CGCGAGGGGCGCGAGACCGCTCTGGAATCGGAATCTCTAAAGCGTTTCTCGTTCGATT 180
Db |||||
Qy 121 CGCGAGGGGCGCGAGACCGCTCTGGAATCGGAATCTCTAAAGCGTTTCTCGTTCGATT 180
Db |||||
Qy 181 CAGCTCCTCTGTGTTTTCGCGCTGGATTTGGCTTAATGGATACTACAGACAACAGAGAAA 240
Db |||||
Qy 181 CAGCTCCTCTGTGTTTTCGCGCTGGATTTGGCTTAATGGATACTACAGACAACAGAGAAA 240
Db |||||
Qy 241 CTTGTTGAAGAGATTTGGCTGCTCTATACAGGAGCACTGAATCAAAAAAATTTGGGGAAG 300
Db |||||
Qy 241 CTTGTTGAAGAGATTTGGCTGCTCTATACAGGAGCACTGAATCAAAAAAATTTGGGGAAG 300
Db |||||
Qy 301 AAATATATCCAAACATGTATAGTACCCCAAAACAATCTCTATCAATATTTGATGAAGATCTTACA 360
Db |||||
Qy 301 AAATATATCCAAACATGTATAGTACCCCAAAACAATCTCTATCAATATTTGATGAAGATCTTACA 360
Db |||||
Qy 361 CAAGTAAATGTGAATCTTAAAGAACTTTAAATTTTCAGGGTTGGGATAAAACATCAATTGGAA 420
Db |||||
Qy 361 CAAGTAAATGTGAATCTTAAAGAACTTTAAATTTTCAGGGTTGGGATAAAACATCAATTGGAA 420
Db |||||
Qy 421 AACACATTTCAATTCATAACACTCTGGGAAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
Db |||||
Qy 421 AACACATTTCAATTCATAACACTCTGGGAAAACAGTGGAAATTAATCTCACTAATGACTACCGT 480
Db |||||
Qy 481 GTACAGCGAGGAGTTTTCAGAAAATGGTGTTTAAAGCAAGCAAGATAAATTTTCACTGGGGA 540
Db |||||
Qy 481 GTACAGCGGAGGAGTTTTCAGAAAATGGTGTTTAAAGCAAGCAAGATAAATTTTCACTGGGGA 540
Db |||||
Qy 541 AAATGCAATATGTCATCTGATGATCAGAGCATAGTTTAGAGGACAAAAATTTCCACTT 600
Db |||||
Qy 541 AAATGCAATATGTCATCTGATGATCAGAGCATAGTTTAGAGGACAAAAATTTCCACTT 600
Db |||||
Qy 601 GAGATGCAAACTACTGCTTTGATGCGGACCGATTTTCAAGTTTGTAGGAAGCAGTCAAA 660
Db |||||
Qy 601 GAGATGCAAACTACTGCTTTGATGCGGACCGATTTTCAAGTTTGTAGGAAGCAGTCAAA 660
Db |||||
Qy 661 GGAAAGGGAAGTTAAAGAGCTTTTATCCATTTTGTGTAGAGTTGGGACAGAGAAAATTTG 720
Db |||||
Qy 661 GGAAAGGGAAGTTAAAGAGCTTTTATCCATTTTGTGTAGAGTTGGGACAGAGAAAATTTG 720
Db |||||
Qy 721 GATTTCAAAGCGATTATTTGATGGAGTGAAGAGTTAGTCTGTTTGGGAAGCAGCTGCT 780
Db |||||
Qy 721 GATTTCAAAGCGATTATTTGATGGAGTGAAGAGTTAGTCTGTTTGGGAAGCAGCTGCT 780
Db |||||
Qy 781 TTAGATCCATTATCATCTCTGCAACCTCTGCAAACTCAACTCAACTCAACTCAACTCAACT 840
Db |||||
Qy 781 TTAGATCCATTATCATCTCTGCAACCTCTGCAAACTCAACTCAACTCAACTCAACTCAACT 840
Db |||||
Qy 841 AATGGCTCATTTGACATCTCTCCCTGCAACAGTGTGACTGGATTTGTTTAAAGAT 900
Db |||||
Qy 841 AATGGCTCATTTGACATCTCTCCCTGCAACAGTGTGACTGGATTTGTTTAAAGAT 900
Db |||||
Qy 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTGTGGAAGTTCTTAAAGAT 960
Db |||||
Qy 901 ACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTTGTGGAAGTTCTTAAAGAT 960
Db |||||
Qy 961 TCTGTTATGTCATGCTGATGAGTCTTATCAAAACAAATTTTCGAGAGCAACAGTACAA 1020
Db |||||
Qy 961 TCTGTTATGTCATGCTGATGAGTCTTATCAAAACAAATTTTCGAGAGCAACAGTACAA 1020
Db |||||
Qy 1021 TTCTCTAGACAGGTGTTTCTCTCATACACTGCGAAAGGAGAGATTCATGAAGCAGTTTGT 1080
Db |||||
Qy 1021 TTCTCTAGACAGGTGTTTCTCTCATACACTGCGAAAGGAGAGATTCATGAAGCAGTTTGT 1080
Db |||||
```

```
Qy 1081 AGTTCAGAACCCAGAAAATGTTTCAGGCTGACCCAGAGAATTATACCAGCCTTCTTGTAC 1140
Db |||||
Qy 1081 AGTTCAGAACCCAGAAAATGTTTCAGGCTGACCCAGAGAATTATACCAGCCTTCTTGTAC 1140
Db |||||
Qy 1141 TGGGAAGACCTCGAGTCGTTTATGATACCATGATGAGAAGTTTGCAGTTTGTACCA 1200
Db |||||
Qy 1141 TGGGAAGACCTCGAGTCGTTTATGATACCATGATGAGAAGTTTGCAGTTTGTACCA 1200
Db |||||
Qy 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGAATTTTGCACAGATGGCTATCAAGACTTG 1260
Db |||||
Qy 1201 CAGTTGGATGGAGAGGACCAACCAAGCATGAATTTTGCACAGATGGCTATCAAGACTTG 1260
Db |||||
Qy 1261 GTAACCTAT 1268
Db |||||
Qy 1261 GGTGCTAT 1268
Db |||||
```

Search completed: August 30, 2005, 19:11:39
Job time : 1921 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 11:31:34 ; Search time 499 Seconds
(without alignments)
10135.734 Million cell updates/sec

Title: US-09-983-000A-1
Perfect score: 3091
Sequence: 1 cacacatgcacgcacgat.....ttatgtagtaagtgtatata 3091

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1263.2	40.9	7941	3	US-09-816-703A-1
2	1113	36.0	6924	1	US-08-015-973-2
3	1113	36.0	6924	2	US-08-448-164-2
4	189	6.1	4338	1	US-08-015-986A-1
5	189	6.1	4338	2	US-08-446-363-1
6	74.8	2.4	1141	4	US-09-806-708B-22
7	55.2	1.8	1141	4	US-09-806-708B-22
8	53.2	1.7	1141	4	US-09-949-016-16536
9	53	1.7	601	4	US-09-949-016-30530
10	53	1.7	601	4	US-09-949-016-37149
11	53	1.7	601	4	US-09-949-016-37163
12	53	1.7	601	4	US-09-949-016-145867
13	53	1.7	601	4	US-09-949-016-146135
14	53	1.7	601	4	US-09-949-016-146403
15	53	1.7	205044	4	US-09-949-016-15851
16	53	1.7	205044	4	US-09-949-016-15852
17	53	1.7	205044	4	US-09-949-016-15853
18	53	1.7	223471	4	US-09-949-016-12387
19	53	1.7	223471	4	US-09-949-016-12724
20	53	1.7	223471	4	US-09-949-016-12725
21	52.8	1.7	601	4	US-09-949-016-30531
22	52.8	1.7	601	4	US-09-949-016-37150
23	52.8	1.7	601	4	US-09-949-016-37164
24	52.8	1.7	601	4	US-09-949-016-145868
25	52.8	1.7	601	4	US-09-949-016-146136
26	52.8	1.7	601	4	US-09-949-016-146404
27	51.8	1.7	7218	1	US-08-232-463-14

28	48	1.6	19124	2	US-08-487-826B-13	Sequence 13, Appl
29	47.6	1.5	137226	4	US-09-949-016-13763	Sequence 13763, A
30	47.6	1.5	212139	4	US-09-949-016-16065	Sequence 16065, A
31	47.4	1.5	251672	4	US-09-949-016-17296	Sequence 17296, A
32	47.4	1.5	251682	4	US-09-949-016-11973	Sequence 11973, A
33	47.4	1.5	263693	4	US-09-949-016-12386	Sequence 12386, A
34	47.4	1.5	263694	4	US-09-949-016-16915	Sequence 16915, A
35	47.2	1.5	17082	4	US-09-949-016-14893	Sequence 14893, A
36	47.2	1.5	107827	4	US-09-949-016-15790	Sequence 15790, A
37	47	1.5	50217	4	US-09-949-016-16067	Sequence 16067, A
38	46.6	1.5	63183	4	US-09-949-016-13047	Sequence 13047, A
39	46.6	1.5	63183	4	US-09-949-016-13048	Sequence 13048, A
40	46.6	1.5	640681	4	US-09-790-988-1	Sequence 1, Appl
41	46.4	1.5	601	4	US-09-949-016-185649	Sequence 185649, A
42	46.4	1.5	601	4	US-09-949-016-185650	Sequence 185650, A
43	46.4	1.5	113966	4	US-09-949-016-12277	Sequence 12277, A
44	46.4	1.5	113967	4	US-09-949-016-17051	Sequence 17051, A
45	46.2	1.5	3095	6	5231168-1	Patent No. 5231168

ALIGNMENTS

RESULT 1
US-09-816-703A-1
; Sequence 1, Application US/09816703A
; Patent No. 6455026
; GENERAL INFORMATION:
; APPLICANT: Melcher, Thorsten
; APPLICANT: Mueller, Sabine
; APPLICANT: Chin, Daniel
; TITLE OF INVENTION: Use of Protein Tyrosine Phosphatase Zeta as a Biomolecular Target
; TITLE OF INVENTION: Treatment and Visualization of Brain Tumors
; FILE REFERENCE: 262/235 AGY
; CURRENT APPLICATION NUMBER: US/09/816.703A
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 7941
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (148)..(7092)
US-09-816-703A-1

Query Match	40.9%	Score 1263.2;	DB 3;	Length 7941;
Best Local Similarity	99.8%	Pred. No. 3.4e-299;		
Matches 1265;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	1	CACACATACGACGACGATCTCACTTCGATCTATACACTGGAGGATTAACAAACAAA	60	
Db	1	CACACATACGACGACGATCTCACTTCGATCTATACACTGGAGGATTAACAAACAAA	60	
Qy	61	CAAAAAAACAATTCCTTCGCTCCCTCCCTCTCCACTCTGAGAAGCAGAGGAGCCGCA	120	
Db	61	CAAAAAAACAATTCCTTCGCTCCCTCCCTCTCCACTCTGAGAAGCAGAGGAGCCGCA	120	
Qy	121	CGGCGAGGGGCGCAGACCGCTCTGGAATGCAATCTCTAAAGCGTTCTCGCTTGCATT	180	
Db	121	CGGCGAGGGGCGCAGACCGCTCTGGAATGCAATCTCTAAAGCGTTCTCGCTTGCATT	180	
Qy	181	CAGCTCTCTGTGTTTTCGCGCTGGATTTGGGCTAATGGATCTACAGACAGAGAAA	240	
Db	181	CAGCTCTCTGTGTTTTCGCGCTGGATTTGGGCTAATGGATCTACAGACAGAGAAA	240	
Qy	241	CTTGTTCAAGAGATTTGGCTGCTCTATACAGGAGCACTGAATCAAAAAATTTGGGGAAG	300	
Db	241	CTTGTTCAAGAGATTTGGCTGCTCTATACAGGAGCACTGAATCAAAAAATTTGGGGAAG	300	
Qy	301	AAATATCCACATGTAATAGCCCAACAAATCTCTTCAATATTGATGAAGATCTTACA	360	

301	AAATATCCAACTGTAATAGCCCAAAACAATCTCCTATCAATAATTGTGAAGAATCTTTACA	360
Db		
361	CAAAGTAAATGTGAATCTTTAAGAAAATTAAAAATTCAGGGTTGGGATAAAAAATCATATTGGAA	420
Qy		
361	CAAAGTAAATGTGAATCTTTAAGAAAATTTAAATTTTCAGGGTTGGGATAAAAAATCATATTGGAA	420
Db		
421	AACACATTCAATTCATAACACTCGGGAAAAACAGTGGAAAAATTAATCTCACATAATGACTACCGT	480
Qy		
421	ARACACATTCAATTCATAACACTCGGGAAAAACAGTGGAAAAATTAATCTCACATAATGACTACCGT	480
Db		
481	GTCAGCGGAGGAGTTTTCCAGAAAATGGTGTTTAAAGCAAGCAAGATAAACCTTTTTCACCTGGGGGA	540
Qy		
481	GTCAGCGGAGGAGTTTTCCAGAAAATGGTGTTTAAAGCAAGCAAGATAAACCTTTTTCACCTGGGGGA	540
Db		
541	AAATGCAATATGTCATCTGATGGATCAGAGCATATAGTTTAGAAGGACAAAAATTTCCACTT	600
Qy		
541	AAATGCAATATGTCATCTGATGGATCAGAGCATATAGTTTAGAAGGACAAAAATTTCCACTT	600
Db		
601	GAGATGCCAAATCTACTGCTTTTGATGCCGACCGATTTTCAAGTTTGTAGAGGACAGAGTCAAA	660
Qy		
601	GAGATGCCAAATCTACTGCTTTTGATGCCGACCGATTTTCAAGTTTGTAGAGGACAGAGTCAAA	660
Db		
661	GGAAAGGGAAGTTTAAGAGCTTTTATCCATTTTGTTCGAGTTGGGCACAGAAGAAAAATTTG	720
Qy		
661	GGAAAGGGAAGTTTAAGAGCTTTTATCCATTTTGTTCGAGTTGGGCACAGAAGAAAAATTTG	720
Db		
721	GATTTCAAAGCGATTAATTGATGGAGTCCGAAGTGTTAGTCGTTTGGGAAGCAGCGTCT	780
Qy		
721	GATTTCAAAGCGATTAATTGATGGAGTCCGAAGTGTTAGTCGTTTGGGAAGCAGCGTCT	780
Db		
781	TTAAGATCCATTCATCTGTTTGAACCTTCTGCCAAAACCTCAACTGACAAAGTATTACATTTC	840
Qy		
781	TTAAGATCCATTCATCTGTTTGAACCTTCTGCCAAAACCTCAACTGACAAAGTATTACATTTC	840
Db		
841	AATGGCTCATTTGACATCTCTCCCTGCACACACACAGTTGACTCGATTGTTTTTAAAGAT	900
Qy		
841	AATGGCTCATTTGACATCTCTCCCTGCACACACACAGTTGACTCGATTGTTTTTAAAGAT	900
Db		
901	ACAATTAGCATCTCTGAAGCCAGTTGGCTGTTTTTTGTGAAGTTCTTACAATGCAACAA	960
Qy		
901	ACAATTAGCATCTCTGAAGCCAGTTGGCTGTTTTTTGTGAAGTTCTTACAATGCAACAA	960
Db		
961	TCTGGTTATGTCATGCTGATGGACTACTTACAAACAAATTTTCGAGAGCACAGTACAAG	1020
Qy		
961	TCTGGTTATGTCATGCTGATGGACTACTTACAAACAAATTTTCGAGAGCACAGTACAAG	1020
Db		
1021	TTCTCTAGACAGGTGTTTTCTCTCATACACTCGGAAAGGAAGATTCATGAAGCAGTTTGT	1080
Qy		
1021	TTCTCTAGACAGGTGTTTTCTCTCATACACTCGGAAAGGAAGATTCATGAAGCAGTTTGT	1080
Db		
1081	AGTTTCAAGAACAGAAAATGTTTCAGGCTGACCCAGAGAATTAATACAGCCTTCTTGTTTACA	1140
Qy		
1081	AGTTTCAAGAACAGAAAATGTTTCAGGCTGACCCAGAGAATTAATACAGCCTTCTTGTTTACA	1140
Db		
1141	TGGGAAGACCTCGAGTCGTTTTATGATACCATGATTCGAGAGTTTGCAGTTTGTGTCACG	1200
Qy		
1141	TGGGAAGACCTCGAGTCGTTTTATGATACCATGATTCGAGAGTTTGCAGTTTGTGTCACG	1200
Db		
1201	CAGTTGGATGGAGGAGCCAAACCAACAGCATGAATTTTTTGACAGATGGCTATCAAGACTTG	1260
Qy		
1201	CAGTTGGATGGAGGAGCCAAACCAACAGCATGAATTTTTTGACAGATGGCTATCAAGACTTG	1260
Db		
1261	GTAACTAT 1268	
Qy		
1261	GGTGCTAT 1268	
Db		

TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
 TITLE OF INVENTION: PHOSPHATASE-BETA
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PENNIE & EDMONDS
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/015,973
 FILING DATE: 10-FEB-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Mistrock, S. Leslie
 REGISTRATION NUMBER: 18,972
 REFERENCE/DOCKET NUMBER: 7683-021
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)790-9090
 TELEFAX: (212)869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6924 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..6924
 US-08-015-973-2

Query Match 36.0%; Score 1113; DB 1; Length 6924;
Best Local Similarity 99.6%; Pred. No. 2.1e-262;
Matches 1116; Conservative 0; Mismatches 5; Indels 0; Gaps 0

Qy	148	ATGCGAATCC	TAAGCGTTTCCTCGCTTGCAAT	CAGTCCCTCTGTGTTTGGCGCCTGAT	207
Db	1	ATGCGAATCC	TAAGCGTTTCCTCGCTTGCAAT	CAGTCCCTCTGTGTTTGGCGCCTGAT	60
Qy	208	TGGGCTAAT	TGGATACTACAGACACAGAGAAA	CTTGTTGAAGAGATTTGGCTGGTCCCTAT	267
Db	61	TGGGCTAAT	TGGATACTACAGACACAGAGAAA	CTTGTTGAAGAGATTTGGCTGGTCCCTAT	120
Qy	268	ACGGAGCAGT	GAATCAAAAAATTTGGGGAAGAAA	TATCCAAATGTAAATAGGCCCAAAA	327
Db	121	ACGGAGCAGT	GAATCAAAAAATTTGGGGAAGAAA	TATCCAAATGTAAATAGGCCCAAAA	180
Qy	328	CAATCTCCT	ATCAATTTGATGAAGATCTTACACAA	GATAAATGTCAATCTTAAAGAAACTTT	387
Db	181	CAATCTCCT	ATCAATTTGATGAAGATCTTACACAA	GATAAATGTCAATCTTAAAGAAACTTT	240
Qy	388	AAATTTTCA	GCGTTGGGATAAAACATCAATTG	GAAAACACATTTCAATTCATAACACTGGGAAA	447
Db	241	AAATTTTCA	GCGTTGGGATAAAACATCAATTG	GAAAACACATTTCAATTCATAACACTGGGAAA	300
Qy	448	ACAGTGGAA	TATTAATCTCACTAATGA	CTACCGTGTCCGCGAGGAGTTTTCAGAAATGGTG	507
Db	301	ACAGTGGAA	TATTAATCTCACTAATGA	CTACCGTGTCCGCGAGGAGTTTTCAGAAATGGTG	360
Qy	508	TTTTAAAG	CAAGCAAGATAACTTTTTTCACTGGG	GGAAAAATGCAATATGTCTCATCTGATGGATCA	567
Db	361	TTTTAAAG	CAAGCAAGATAACTTTTTTCACTGGG	GGAAAAATGCAATATGTCTCATCTGATGGATCA	420
Qy	568	GAGCATAGT	TTTAGAAGCAAAAAATTTCCACTTG	GAGATGCAAAATCTACTGCTTTTGATCGG	627

RESULT 2
US-08-015-973-2
; Sequence 2, Application US/08015973
; Patent No. 5604094
; GENERAL INFORMATION:
; APPLICANT: Schlössinger, Joseph

Db 421 GAGCAGAGTTTGAAGGACAAAAATTTCCATCTTGAGATCAAATCTACTGCTTTGATGCA 480
Qy 628 GACCGATTTTCAAGTTTTCAGGAAGCAGTCAAAGGAAAAGGGAAGTTAAGAGCTTTATCC 687
Db 481 GACCGATTTTCAAGTTTTCAGGAAGCAGTCAAAGGAAAAGGGAAGTTAAGAGCTTTATCC 540
Qy 688 ATTTCGTTGAGTTGGGACAGAGAAAATTTGGATTTCAAAGCGGATTTATGATGAGTC 747
Db 541 ATTTCGTTGAGTTGGGACAGAGAAAATTTGGATTTCAAAGCGGATTTATGATGAGTC 600
Qy 748 GAAAGTTTGTAGTCTTTTCGGAGCAGGCTGCTTTAGATCCATTCATCTGTTGACCTT 807
Db 601 GAAAGTTTGTAGTCTTTTCGGAGCAGGCTGCTTTAGATCCATTCATCTGTTGACCTT 660
Qy 808 CTGCCAACTCAACTGACAGATTAATACATTTACAATGGCTCAATGACATCTCTCCCTGC 867
Db 661 CTGCCAACTCAACTGACAGATTAATACATTTACAATGGCTCAATGACATCTCTCCCTGC 720
Qy 868 ACAGACACAGTTGACCTGGATTTGTTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
Db 721 ACAGACACAGTTGACCTGGATTTGTTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 780
Qy 928 GCTGTTTTTGTGAGTTCTTCAATGCAACCAATCTGTTATGTCTATGCTGATGACTAC 987
Db 781 GCTGTTTTTGTGAGTTCTTCAATGCAACCAATCTGTTATGTCTATGCTGATGACTAC 840
Qy 988 TTCAAAAACAATTTTCGAGGACCAACAGTACAAAGTTCTCTAGACAGGTGTTTTTCCTCATAC 1047
Db 841 TTCAAAAACAATTTTCGAGGACCAACAGTACAAAGTTCTCTAGACAGGTGTTTTTCCTCATAC 900
Qy 1048 ACTGGAAGGAGAGATTCATGACAGCTTTGTTAGTTTCAGAAACCAAGAAATGTTTCAGGCT 1107
Db 901 ACTGGAAGGAGAGATTCATGACAGCTTTGTTAGTTTCAGAAACCAAGAAATGTTTCAGGCT 960
Qy 1108 GACCCAGAGATTTATACAGGCTCTTCTGTTATACAGGAAAGCTCGAGTCGTTTATGAT 1167
Db 961 GACCCAGAGATTTATACAGGCTCTTCTGTTATACAGGAAAGCTCGAGTCGTTTATGAT 1020
Qy 1168 ACCATGATTGAGAGTTTGCAGTTTGTACAGCAGTTTGGATGGAGAGACCAACCAAG 1227
Db 1021 ACCATGATTGAGAGTTTGCAGTTTGTACAGCAGTTTGGATGGAGAGACCAACCAAG 1080
Qy 1228 CATGAATTTTGCAGATGGCTTATCAAGACTTTGGTAACTAT 1268
Db 1081 CATGAATTTTGCAGATGGCTTATCAAGACTTTGGGTGCTAT 1121

RESULT 3

US-08-448-164-2 :
; Sequence 2, Application US/08448164
; Patent No. 5925536
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-BETA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSES: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,164
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/015,973
; FILING DATE: 10-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6924 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..6924
; US-08-448-164-2

Query Match 36.0%; Score 1113; DB 2; Length 6924;
Best Local Similarity 99.6%; Pred. No. 2.1e-262;
Matches 1116; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 148 ATGCGAATCCTAAAGCGTTTCCTCGCTTGCAATTCAGCTCCTCTGTTGCGGCTGGAT 207
Db 1 ATGCGAATCCTAAAGCGTTTCCTCGCTTGCAATTCAGCTCCTCTGTTGCGGCTGGAT 60
Qy 208 TGGGCTTAATGGATACTACAGACAAACAGAGAAAACTTGTGTAAGAGATTTGGCTGCTTAT 267
Db 61 TGGGCTTAATGGATACTACAGACAAACAGAGAAAACTTGTGTAAGAGATTTGGCTGCTTAT 120
Qy 268 ACAGAGACCTGAATCAAAAAAATTTGGGAAAAGAAAATATCCAAATGTATAGCCAAAA 327
Db 121 ACAGAGACCTGAATCAAAAAAATTTGGGAAAAGAAAATATCCAAATGTATAGCCAAAA 180
Qy 328 CAATCTCTTCAATATTTGATGAAGATCTTACACAGTAAATGTGAATCTTTAAGAACTT 387
Db 181 CAATCTCTTCAATATTTGATGAAGATCTTACACAGTAAATGTGAATCTTTAAGAACTT 240
Qy 388 AAATTTTCAGGTTGGGATAAAACATCATTTGAAAAACACATTCATTCATAAACCTGGGAAA 447
Db 241 AAATTTTCAGGTTGGGATAAAACATCATTTGAAAAACACATTCATTCATAAACCTGGGAAA 300
Qy 448 ACAGTGAATTAATCTCACTAATGACTACCGTGTCCAGCGGAGAGTTTCAGAAATGGTG 507
Db 301 ACAGTGAATTAATCTCACTAATGACTACCGTGTCCAGCGGAGAGTTTCAGAAATGGTG 360
Qy 508 TTTAAAGCAAGCAGATAACTTTTTCAGTGGGAAAATGCAATATGTCATCTGATGGATCA 567
Db 361 TTTAAAGCAAGCAGATAACTTTTTCAGTGGGAAAATGCAATATGTCATCTGATGGATCA 420
Qy 568 GAGCATAGTTTGAAGAGCAAAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATGCG 627
Db 421 GAGCACAGTTTGAAGAGCAAAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATGCA 480
Qy 628 GACCGATTTTCAAGTTTGGAGAGCAGTCAAGGAAAAGGAAAGTTAAGAGCTTTATCC 687
Db 481 GACCGATTTTCAAGTTTGGAGAGCAGTCAAGGAAAAGGAAAGTTAAGAGCTTTATCC 540
Qy 688 ATTTTGTTCAGGTTGGGACAGAGAAAATTTGGATTTCAAAGCGATTTATGATGGAGTC 747
Db 541 ATTTTGTTCAGGTTGGGACAGAGAAAATTTGGATTTCAAAGCGATTTATGATGGAGTC 600
Qy 748 GAAAGTGTAGTCTTTTTCGGAAAGCAGGCTGCTTTTAGATCCATTCATCTGTTGAACTT 807
Db 601 GAAAGTGTAGTCTTTTTCGGAAAGCAGGCTGCTTTTAGATCCATTCATCTGTTGAACTT 660
Qy 808 CTGCCAACTCAACTGACAGATTAATACATTTAAGATGGCTCATTTGACATCTCTCCCTGC 867
Db 661 CTGCCAACTCAACTGACAGATTAATACATTTAAGATGGCTCATTTGACATCTCTCCCTGC 720

QY 868 ACAGACACAGTTGACTGATGTTTAAAGATACAGTTAGCATCTCTGAAGCCAGTTG 927
Db |||||
QY 721 ACAGACACAGTTGACTGATGTTTAAAGATACAGTTAGCATCTCTGAAGCCAGTTG 780
Db |||||
QY 928 GCTGTTTTTGTGAAGTTCTTACAAATCAATCTGGTTATGTCATGCTGATGACTAC 987
Db |||||
QY 781 GCTGTTTTTGTGAAGTTCTTCAATATGCAACATCTGGTTATGTCATGCTGATGACTAC 840
Db |||||
QY 988 TTAACAAAACATTTTCGAGAGCAACAGTCAAGTTCTCTAGACAGGTGTTTTCTCATAC 1047
Db |||||
QY 841 TTAACAAAACATTTTCGAGAGCAACAGTCAAGTTCTCTAGACAGGTGTTTTCTCATAC 900
Db |||||
QY 1048 ACTGGAAGGAGAGATTCTAGACAGTTTGTAGTTTCTAGACACCAAGAAATGTTGAGCT 1107
Db |||||
QY 901 ACTGGAAGGAGAGATTCTAGACAGTTTGTAGTTTCTAGACACCAAGAAATGTTGAGCT 960
Db |||||
QY 1108 GACCCAGAGAAATTATACAGAGCTTCTGTATCATGGGAAAGACCTCGAGTGTGTTTATGAT 1167
Db |||||
QY 961 GACCCAGAGAAATTATACAGAGCTTCTGTATCATGGGAAAGACCTCGAGTGTGTTTATGAT 1020
Db |||||
QY 1168 ACCATGATGAGAGTTTGCAGTTTGTATACAGAGTTTGGATGGAGAGCAACCAAG 1227
Db |||||
QY 1021 ACCATGATGAGAGTTTGCAGTTTGTATACAGAGTTTGGATGGAGAGCAACCAAG 1080
Db |||||
QY 1228 CATGAATTTTGCAGAGTGGCTATCAGACTTGTGTAACAT 1268
Db |||||
QY 1081 CATGAATTTTGCAGAGTGGCTATCAGACTTGTGTAACAT 1121
Db |||||

RESULT 4

US-08-015-986A-1
; Sequence 1, Application US/08015986A
; Patent No. 5532123
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-GAMMA
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESS: PENNIE & EDMONDS
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/015,986A
; FILING DATE: 10-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/POCKET NUMBER: 7683-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4338 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4335

US-08-015-986A-1

Query Match 6.1%; Score 189; DB 1; Length 4338;
Best Local Similarity 54.0%; Pred. No. 5e-36;
Matches 409; Conservative 0; Mismatches 345; Indels 3; Gaps 1;
QY 258 CTGTCCTATACAGGAGCAGTGAATCAAAAATTTGGGGAGAAATATCCAAATGTA 317
Db |||||
QY 177 CTGGGCTTACTCTGGTCCCTATGCTCTAGACATCTGGTCTAGTGTGCTGCTG 236
Db |||||
QY 318 TAGCCCAAAAACATCTCTATCAATATTTGATGAAGATCTTACACAAGTAAATGTGAATCT 377
Db |||||
QY 237 GAGCCGTCAACAGTCTCTATTTAGACATTTTAGACCATGATGCGGTGTTGGGGAAGAATA 296
Db |||||
QY 378 TAGAAACATTAATTTTCAGGTTGGGATAAAAATCAATCTTTGGAAAAACAATTCATTCAATTA 437
Db |||||
QY 297 CCAGGAATCGCAATCGATGGCTTCGACAATGAGTCTTTTAAACAAAACCTGGATGA 356
Db |||||
QY 438 CACTGGGAAAAACAGTGGAAATTAATCTCACTAATGACTACCGTGTCTAGCGGAGGAGTTTC 497
Db |||||
QY 357 CACAGGAAAAACAGTCCCATCTCTTCTGAAGAGACTATTTTGTCTAGTGGAGCTGTCT 416
Db |||||
QY 498 AGAAATGCTGTTTAAAGCAAGCAAGATACTTTTCTCTGGGGAAAAATGCAATATGTCTATC 557
Db |||||
QY 417 ACCTGGCAGATTCAAAGCTGAGAAGGTGGAAATTTCTCTGGGGCCACAGCAATGSGCTCA 474
Db |||||
QY 558 TGATGGATCAGAGCATGATTTAGAGGACAAAAATTTTCCACTTGGAGATGCAATCTACTG 617
Db |||||
QY 475 -GCGGGCTCTGAACACACAGCATCAATGGCAGGAGGTTTCTCTGTTGAGATGCAGATTTTCTT 533
Db |||||
QY 618 CTTTGTATGCGGACCGGATTTTCAAGTTTGTAGGAAGCAGTCAAAAGGAAAAAGGAGTTTAAG 677
Db |||||
QY 534 TTACATCCAGATGACTTTGACAGCTTTTCAACCGCAATTTCTGAGAACAGAAATATCGG 593
Db |||||
QY 678 AGCTTTATCCATTTTGTGAGGTGGGACAGAGAAATTTTGGATTTCAAAGCGATTAT 737
Db |||||
QY 594 AGCCATGGCCATATTTTCAAGTCACTCCGAGGACAAATTTCTGCACTGGATCCTATTAT 653
Db |||||
QY 738 TGATGGAGTCGAAAGTGTAGTGTGTTTGGGAGCAGGCTGCTTTAGATCCATTCATCT 797
Db |||||
QY 654 CCACGGGTTGAAGGGTGTCTGATCATGAGAAGGAGACCTTTCTGGATCCTTTCTGTCCT 713
Db |||||
QY 798 GTTGAACCTTCTGCCAAACTCACTGACAAAGTATTACATTTTACAAATGGCTCATTTGACATC 857
Db |||||
QY 714 CCGGGACCTCTGCTGCATCCCTGGGCGAGCTATTATCGGTACACAGGTTCTCTTGACAC 773
Db |||||
QY 858 TCCTCCCTGCACAGACACAGTTGATGCTGGATTTGTTTAAAGATACAGTTAGCATCTCTGA 917
Db |||||
QY 774 ACCACCGTGTAGCGAAATAGTGGAGTGGATAGTCTTCCGGAGACCCGCTCCCATCTCTTA 833
Db |||||
QY 918 AAGCAGTTGGCTGTTTTTTTGTGAGTTCTTACAAATGCAACAATCTGGTTATGTCATGCT 977
Db |||||
QY 834 CCATCAGCTTTGAGGCTTTTATTTCCATCTTCCACCGGAGCAGCAAGCATGTGCAAGTC 893
Db |||||
QY 978 GATGGACTACTTACAAAACAAATTTTCGAGAGCAACAG 1014
Db |||||
QY 894 GTGGAGTATCTGAGAAATAACTTTCCACACAGCAG 930
Db |||||

RESULT 5

US-08-446-363-1
; Sequence 1, Application US/08446363
; Patent No. 5891700
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-GAMMA
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESS: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York

COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/446,363
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/015,986
 FILING DATE: 10-FEB-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Mirock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7683-028
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4338 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..4335
 US-08-446-363-1

Query Match 6.1%; Score 189; DB 2; Length 4338;
 Best Local Similarity 54.0%; Pred. No. 5e-36;
 Matches 409; Conservative 0; Mismatches 345; Indels 3; Gaps 1;

```

QY 258 CTGGTCTCTATACAGGACACTGAATCAAAAAATTTGGGGAAGAAATATCCAAATGTAA 317
DB 177 CTGGGCTACTCTGGTGCTATGGTCTCTGAGCACTGGGTGCTAGTGTGAGTGTGG 236
QY 318 TAGCCCAAAACATCTCTATCAATTAATGATGAAGATCTTACACAAAGTAATGTGAATCT 377
DB 237 GAGCGCTACCAAGTCTCTATTGACATTTAGACAGATGCGGTGTTGGGGAAGAATA 296
QY 378 TAGGAACCTTAATTTTCAGGGTTGGATGAACATCATTTGGAACACATTCATTCTATAA 437
DB 297 CCAGGAACCTGCAACTCGATGGCTTCGACAAATGAGTCTTTAAACAAAACCTGGATGAAA 356
QY 438 CACTGGGAAAACAGTGGAAATTAATCTCACTAATGACTACCGTGTGAGCGGAGGATTTTC 497
DB 357 CACAGGAAAACAGTCGCCATCTCTGAGAGACACTATTTTGTGAGTGGAGCTGGTCT 416
QY 498 AGAAATGGTGTTTAAAGCAAGCAAGATACTTTTCACTGGGAAAATGCAATATGTCATC 557
DB 417 ACCTGGCAGATCAAAAGCTGAGAAGTGGAAATTTCACTGGGGCCACAGCAATGGCTCA 474
QY 558 TGATGATCAGAGCATAGTTTGAAGGACAAAATTTTCCACTTGGAGATGCAATCTACTG 617
DB 475 -GCGGGCTCTGAACACAGCATCAATGGGAGGAGGTTTCTGTGTGAGATGCGAGATTTCTT 533
QY 618 CTTTGCATCGGACCGATTTTCAAGTTTGGAGAACAGTCAAGGAAAAGGAGGATTAAG 677
DB 534 TTCAATTCAGATGACTTTGACAGCTTTCAAAACCGCAATTTCTGAGAACAGAAATATCGG 593
QY 678 AGCTTATTCATTTTGTGTGAGGTTGGGACAGAGAAAATTTGGATTTCAAAGCGATTAAT 737
DB 594 AGCCATGCGCATATTTTTCAGTCAAGTCAAGTCCGAGGAGCAATCTGCACTGGATCTATTAT 653
QY 738 TGATGAGTCGAAAGTGTGTAGTCGTTTGGGAGCAGGCTGCTTTAGATTCATTCATCTACT 797
DB 654 CCACGGGTTGAAGGGTGTGTCATCATCATGAGAGGAGGACCTTTTCTGGATCTCTTTCGTCT 713
  
```

```

QY 798 GTTGAACCTTCTGCGAAACTCAACTGCAAGTATTACATTTACAATGGCTCAATGACATC 857
DB 714 CCGGACCTCTGCTGCTGATCCCTGGGCGAGCTATTATCGGTACACAGGTTCTTTGACCAC 773
QY 858 TCTTCCCTGCGACAGACAGATGAGTGTGATTTTAAAGATACAGTTAGATCTCTGA 917
DB 774 ACCACCGTGTAGCGAAATAGTGGAGTGTAGTCTTCCGGAGACCCGTCCTCTTTA 833
QY 918 AAGCAGTTGGCTGTTTTTGTGAGTCTTACAAATCTTACAAATCTGTTATGTCATGCT 977
DB 834 CCATCAGCTTGAGGCTTTTTTATTCATCTTCCACCGAGCAGCAAGACCATGTCAAGTC 893
QY 978 GATGACTACTTACAAAACAAATTTTCGAGAGCAACAG 1014
DB 894 GGTGGAGTATCTGAGAAATACTTTGACCAACAGCAG 930

RESULT 6
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAEL promoters
US-09-806-708B-22

Query Match 2.4%; Score 74.8; DB 4; Length 1141;
Best Local Similarity 11.5%; Pred. No. 2.7e-08;
Matches 115; Conservative 379; Mismatches 500; Indels 8; Gaps 3;
QY 1026 TAGCAGGTGTTTCTCATACACTGGAAGGAAGAGATTCATGAGCAGTGTGTAGTTC 1085
DB 127 YWRWAMBDTVDDHHVVTAMNNATTCMDKDDKRTTRWMMKNNNATGWDDDTTKYHMNN 186
QY 1086 AGRACCAAGAAATCTTCAGGCTGACCCAGAGAAATATATACCAGCTTCTTGTATCATGGA 1145
DB 187 NGCBTVMVRYKTDROWSKRMNYGMBWKNWSYDVTYYWVWDDMKCKRKVRWVRTRG 246
QY 1146 AAGACTCGAGTCTTTATGATACCATGATTTGAGAGTGTTCAGTGTGTGTACACGAGTT 1205
DB 247 RMRYWVWBTAFHRRYNNNGWTBWA YRWTNNNNNNNAKAMCKRKYWGNRABVNSTC 306
QY 1206 GGATGGAGAGGACCAACCAAGCATGAATTTTTCAGAGAT----GGCATATCAAGACTGG 1261
DB 307 TTWKSRTTKVRTS CWANNCRAGDANKDKHKKWKSAAAGVYNNNNNNNNNTTYYKARHBAW 366
QY 1262 TAACATATATGATGATGTTGTTTATACATAGGGAATCAATATATTAATTTTCAAGGTAAGA 1321
DB 367 DWVHWSAKWHAANAHSRKWTBYRKRTWNNNNNGTTTWMKRWMAWYRMDMDWBGTYN 426
QY 1322 ACTTACAAATGGTGTGTATATTTTCTCCATTAATTTTAGACTTTATGTGAAGGTGGG 1381
DB 427 NNNNGRTTYTGTKKQWTVYKWKANNCKRWADHKTCTHNTTMMKMTKYNNCYWKS 486
QY 1382 GTAGGCTGAGTATTTTAAATTTTAAATAATTTTAAATTTTAAATTTTAAAGCTATATACTAAATAT 1441
DB 487 MTNGKSHRBAAAVVTWYMMWRRYAHANNNNNDYWWKACTWYKYBVCBCKWNNNYAAWYTK 546
  
```

QY 1442 GTTTAAAGTTACATTTAAATTAATGGATATCATTAACCTTTGGCAACAATAACACTATAGAG 1501
DB 547 SSWNTSYRKYKTNNSWRWSRSDTRSMGRANNYARABHYGYKWNTRWBWS-HTWBHBR 605
QY 1502 TAGATACATATGACTTATGAAGCTGAGATCATTTAGTGTGGCTTTCTTAAAGATTTCAGT 1561
DB 606 GAATYMBMBTBAKCHCKWAKYKAKYAGAGSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 665
QY 1562 TGTAGAAATAGTCCAGAAATCTCAGTGCCTGATACATTTTATATGTGTCTTCCATACG 1621
DB 666 SRWYAMANAKEYYKBAANNAAYTHANNWGWGNNATDTRRTMKNNNNNNNNNNNNNNNN 725
QY 1622 CTATATCAGCAGAGAAAGTAGTAGGAGGACATACAAGTCTCTTTGTTGGCCACAAA 1681
DB 726 NNAKNAASAQNYAAAVKAAKGRWANKWAMRGWHAADAAAABTTDKRNNGAYTKYTTNN 785
QY 1682 AATTTTCAGATAACAGCTGGGAAGTCATGATTTGGGTCAAGAACTTTGGGATGTAAGAAA 1741
DB 786 NNTYRGVVTNTAARDGWANN 845
QY 1742 CATTTCTTCAAAAAGATCCACCCCTGCTCCCTCCACGCGATGCGAATAAAGTACA 1801
DB 846 YAWNTKWWYTTDDRBAAYTNNNNNNRMAYYGAYDDYAYVMSDTCDAMKWDATKNN 905
QY 1802 GATTCCTTTGGCTGAGATGTCAGTATTAACCTTTGCTCTGGTAGGGAAGTGTGG 1861
DB 906 ATTYNRGTAWRTNNNNNNMTKTYBYBHAANNNNNNNNNNNNNNNNNNNNNNNNNNNN 962
QY 1862 CCATAGATTAGGGTGTAGTGCACAACTTCATCTGGATGTAGTCCAGAAAGTCCCCAC 1921
DB 963 CWNCTTTCTCKYKNCCTWYTWTTTTRTWYATRWKTNNAATGSMTRCNAWGKNNYWTG 1022
QY 1922 TGCAGTTTAAAGCAGTGGACTCTGCACTCAGGCACCTAGAGTCTCTGCAAGTCTGGGA 1981
DB 1023 WKTRTAVRMATRWKWKWATGSMWNTSNVYARWAYKTRAYKGYYNACAWRWGKAT 1082
QY 1982 ACCTGCATTTAAATAAATGCATTAATAATGTTTCATA 2023
DB 1083 CYMTDANWWTACATSWMATHKYNHMKCNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1124

RESULT 7

US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

Query Match 1.8%; Score 55.2; DB 4; Length 1141;
Best Local Similarity 11.8%; Pred. No. 0.0017;
Matches 125; Conservative 370; Mismatches 547; Indels 15; Gaps 2;
QY 1999 AATGCACATTAATTAATGTTTTCATATCATGTGGCAAAAATGGATAAAATTTTAGTAACCT 2058
DB 1138 WRYTCHGNTTKTKYKKAANNNNNNNGKDWNRMDATKWSATGTAWNTNHAKRGTATMC 1079

QY 2059 TTTAAATTCAGTTCGCTGGAATATATGGACACACAATGACCTGGGAAAAATCGTGAATAAATA 2118
DB 1078 WYWTGTTNRWCMRTYAMRTYWTTRSNANWSCATKBMMWTMKWYATKRTAAYAMWCAWR 1019
QY 2119 GTAATAAAAATGTTTATTTTCATAAATACGTCGAAGAAGATAATCTATTAATCTGTTCTTGA 2178
DB 1018 NNNWCATNGYAKSCATNNAMWYATTTRWAAAYAAKWARWAGNNRMWYGAAGNKGWCMAA 959
QY 2179 TATATATTTGCAAGAAAAAGAGATAAATCTAGTTGTTGTCACCTTTTTCACATTCGCTCTGTT 2238
DB 958 MATWGBWADTAGCMCNNNNNNTDVRMARAKANNNNNNNAYWTACYNRAAATNNKMAATH 899
QY 2239 TGCAAATGCCCCCATTTATTTGCTAAAATAATTAATTTTGTAGTTTGTAGTACTAAATTA 2298
DB 898 WMKWTGHAHSKRTRHRTTCR-----RTKYNNNNNNNARTVYWHHAARRRMMNAW 848
QY 2299 TGAATTTGATGAGTTCTGCGCTAAAGAAATGAAACTTCTCTGAAACTAAATCTGATTTTAAAA 2358
DB 847 TRNN 788
QY 2359 AGCAAAAAAAGAGAGCTAGCTTTCCAGTTCTTCATAAATTCACAAATACCAAGT 2418
DB 787 NNNNAARMARTCNVYHAAVTTHTDWCYKTMNTWYDMWMTTMBTTTTTNNMTTSTNMT 728
QY 2419 TTAATAAGCAACATTCATAAACTTTTCTAGTTTAAATAAATAAGAGTATTTTCCAC 2478
DB 727 NNN 668
QY 2479 GGACCGGAGAAAAAGTTTTCTAGGAAAGATACCTAGTGTGTGGTAGTCTCTATGAGAA 2538
DB 667 YSTRHHYTGATNN 608
QY 2539 TAACATTTGATATAATTAACATCTTTCTTTTAGGTGCTATTTCTCAATAATTTGCTAC 2598
DB 607 TCTVVDVSDVWVWYVW 548
QY 2599 CCAATATGAGTATGTTCTTCAGATAGTACCATATGACCTAATGCTTATATGAGAAAT 2658
DB 547 SMARWTTTRNN 488
QY 2659 ACAGCGACCAACTGATTGTCGACATGCT-----ACTGATAATCTCTGTAAGTGCACCC 2712
DB 487 KSMWRGNNNRAMKMWAAANDAGMDHWTYMGNNNTMMRRAMKMMNNMNCRRAYCCNNN 428
QY 2713 AGATACATCATATATTAATCAATAAATGAGGTAGTTTAAATCTGATGATGATGATG 2772
DB 427 NNRAVCVWHKHWWRWTKYWKAAACNNNNNBKAMTRVAMWYSRDTTNTDMMWTSDBW 368
QY 2773 CTTTCTCTCATATTTCTTTGGCCAAAAGGCAAGTATTTCTCTTAAGTCTGATGTC 2832
DB 367 HWYTVDTMMRAMNN 308
QY 2833 CGGTAATTTTGGGCGATGGGACCCATTTCTCATTCAGCAGGTCTGGTGCCACACAAT 2892
DB 307 AGASNBVTYKWRMTYMGKTMTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 248
QY 2893 AAGTAAACTTATCTTAATATGAGTTTACCATTGTTGTAATAAATAAGAGTGAATAACA 2952
DB 247 YCAYBWWYBMYGKHBBWRRABHRSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 188
QY 2953 TTTATAACATTTGTAATAATCATTAATAAGAAATGCTATGTAATGTTGAGACGTGTTATT 3012
DB 187 NNNKKDDMAHHWCATNN 128
QY 3013 TTGATAATTAAGAGTTGGTTTAAATTTGTTATTATT 3049
DB 127 RYDYWCAMCWNNAKAKVRTAMKHWYTTDRVVSANNT 91

RESULT 8

US-09-949-016-16536
; Sequence 16536, Application US/09949016

```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16536
; LENGTH: 114139
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)..(114139)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16536
```

```
Query Match 1.7%; Score 53.2; DB 4; Length 114139;
Best Local Similarity 43.0%; Pred. No. 0.043;
Matches 226; Conservative 0; Mismatches 297; Indels 2; Gaps 1;

QY 2110 AAATAAATAGTATAAATAAGTTTATTTTCATAAATTACGTGAGAGAGATAAATCTTACT 2169
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83352 AAATAAATATATAAATATATATATATATATATATATATATATATATATATATATAT 83411
QY 2170 GTTCTTGCAATATATATGTGCAAGAAAAGAGATAACTAGTTGTTTCACATG 2229
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83412 TATATATACATAAATAATATATAAATAATATATATATATATATATATATATATATAA 83471
QY 2230 CTCCTGTTGTCAAATGCCCCCAATTTATTTCTCTAAATATTAATTTTGTAGTTAGT 2289
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83472 ATATATATATATATATACNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 83531
QY 2290 ACTAATTTATGAATTTGATGAGTCTCGCTAAAAATGAAACTTCTGAAACTTAAATCTGA 2349
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83532 ATAAATATATATATATACATAAATATATAAATATATATATATATATATATATATATAA 83591
QY 2350 TTTTAAAAGCAAAAAAAGCCTAGCTTTCCAGTCTTCCATATATTCACAAT 2409
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83592 ATATATATATATATACATAAATATATATATATATATATATATATATATATATATATAT 83649
QY 2410 ACCACAAGTTTAACCTAAGCAACATTCGATCAAACTTTCTTAGGTTAATAAATAGAGT 2469
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83650 ATAAATATATATATATACATAAATATATAAATATATATATATATATATATATATATAT 83709
QY 2470 ATTTCCAGGACCGAGGAGAAAAGTTTCTTAGGAAAGATACCTAGTGTGTGTTGATC 2529
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83710 ATATAAATATATATATATACATAAATATATAAATATATATATATATATATATATATAT 83769
QY 2530 CTATGAGATAACATTTGTATATATCTCAACATCTTCTTTAGGGTGTCTATCTCAATA 2589
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83770 ATATAAATAATATATATATATATATATATATATATATATATATATATATATATATAT 83829
QY 2590 ATTTGCTACCAATATGATGTTATGTTCTTCAGATAGTAGCCATAT 2634
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83830 ATATAAATAAATATATATATATATATATATATATATATATATATATATATATATATAT 83874
```

```
RESULT 9
US-09-949-016-30530
; Sequence 30530, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30530
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-30530

Query Match 1.7%; Score 53; DB 4; Length 601;
Best Local Similarity 46.8%; Pred. No. 0.0044;
Matches 167; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 1233 ATTTTGACAGATCGCTATCAAGACTTGTAACTATATGATCAGTTGTTTACATAGGTT 1292
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
187 ATAATTTATATATTTTAAATATATATAAATATATATTTTATATATATATATATTTTAA 246
QY 1293 AACATTATAATTAATTTCCAGGTGAAGACTTCAAAATGTTGTATATATTTTCTCTCC 1352
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
247 ATATATATAAATATATTTTATATATATATATATTTTAAATATATATAAATATATTTTAT 306
QY 1353 ATTTACTTTTACACTTTTATGTAAGTGGGAGGCTGAGTATTTTAAATTTTAAAAAAA 1412
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
307 TTTATATTTTAAATATATATAAATATATTTTATATATTTTAAATATATATAAAT 366
QY 1413 ATTTTAAATAGAGCTATACATAAATATGTTTAAAGTTACATTTAAATTAATGGATATC 1472
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
367 ATAATTTATATATTTTATTTTAAATATATATAAATATATTTTATATATTTTATATTTA 426
QY 1473 ATAACTTTGCCAACATAAACAATAGATAGATAGATACATATGACTTATGAACGAGATCA 1532
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
427 AAATATATATAAATATATTTTATATTTTAAATATATTTGATAAATTTTGATGATATAACCT 486
QY 1533 TTTAGTGTGGCTTTCTTTAGATTTTCAGTTGCTAGAAATCTCCAGATCTCCAGTGCC 1589
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
487 TATAATGATATAAATTTTAAATATATTTTGTGATTTATTTTCTCTCTCTATAGTGCC 543
```

```
RESULT 10
US-09-949-016-37149
; Sequence 37149, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37149
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-37149
```

Query Match 1.7%; Score 53; DB 4; Length 601;
Best Local Similarity 46.8%; Pred. No. 0.0044;
Matches 167; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 1233 ATTTTGCAGATGGCTATCAAGACTTGGTAACACTATATGATCAGTTGTTTACATAGGGT 1292
DB 187 ATATTTTATATATTTTAAATATATATAAATATATATTTTATATATTTTATATTTTAAA 246

QY 1293 AACATTATAATTAATTTCCAGGTAAGAACTTACAAATGGTTGTATATATTTTCTCC 1352
DB 247 ATATATATAAATATATTTTATATATTTTAAATATATATAAATATATATTTTATATA 306

QY 1353 ATTACTTTTACACTTTATGTGAAGTGGGTAGGCTGAGTATTTTAAATTTAAATTTAAA 1412
DB 307 TTTATATTTAAATATATATAAATATATTTTATATATTTTAAATATATATAAAT 366

QY 1413 ATTTTAAATTAAGACTATACATAATTAATGTTTAAAGTTACATTTAAATGAATGCATC 1472
DB 367 ATATTTTATATATTTTAAATATATATAAATATATTTTATATATTTTATATTTTAA 426

QY 1473 ATAACTTTGCCAACATAAACAATAGAGTAGATACATATGACTTATGAATCGGAGATCA 1532
DB 427 AAATATATAAATATATTTTATATTTTAAATATATTTGATAAATTTTGATGATATAACCTAA 486

QY 1533 TTTAGTGTGGCCTTTCTTAAAGATTTCAGTTCTGTAAGTAGTCCAGAACTCTCAGTGCC 1589
DB 487 TATAATGTAATAAATTTTAAATATATTTTGAATTTTATTTTGTCTCTTCTATAGTGCC 543

RESULT 11
US-09-949-016-37163
; Sequence 37163, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37163
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-37163

Query Match 1.7%; Score 53; DB 4; Length 601;
Best Local Similarity 46.8%; Pred. No. 0.0044;
Matches 167; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 1233 ATTTTGCAGATGGCTATCAAGACTTGGTAACACTATATGATCAGTTGTTTACATAGGGT 1292
DB 187 ATATTTTATATATTTTAAATATATATAAATATATTTTATATATTTTATATTTTAAA 246

QY 1293 AACATTATAATTAATTTCCAGGTAAGAACTTACAAATGGTTGTATATATTTTCTCC 1352
DB 247 ATATATAAATATATTTTATATATTTTAAATATATATAAATATATATTTTATATA 306

QY 1353 ATTACTTTTACACTTTATGTGAAGTGGGTAGGCTGAGTATTTTAAATTTAAATTTAAA 1412
DB 307 TTTATATTTAAATATATATAAATATATTTTATATATTTTAAATATATATAAAT 366

QY 1413 ATTTTAAATTAAGACTATACATAATTAATGTTTAAAGTTACATTTAAATGAATGCATC 1472
DB 367 ATATTTTATATATTTTAAATATATATAAATATATTTTATATATTTTATATTTTAA 426

QY 1473 ATAACTTTGCCAACATAAACAATAGATAGATACATATGACTTATGAATCGGAGATCA 1532
DB 427 AAATATATAAATATATATTTTATATTTTAAATATATTTGATAAATTTTGATGATATAACCTAA 486

QY 1533 TTTAGTGTGGCCTTTCTTAAAGATTTCAGTTCTGTAAGTAGTCCAGAACTCTCAGTGCC 1589
DB 487 TATAATGTAATAAATTTTAAATATATTTTGAATTTTATTTTGTCTCTTCTATAGTGCC 543

RESULT 12
US-09-949-016-145867
; Sequence 145867, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145867
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145867

Query Match 1.7%; Score 53; DB 4; Length 601;
Best Local Similarity 46.8%; Pred. No. 0.0044;
Matches 167; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 1233 ATTTTGCAGATGGCTATCAAGACTTGGTAACACTATATGATCAGTTGTTTACATAGGGT 1292
DB 187 ATATTTTATATATTTTAAATATATATAAATATATTTTATATATTTTATATTTTAAA 246

QY 1293 AACATTATAATTTAAATTTCCAGGTAAGAACTTACAAATGGTTGTATATATTTTCTCC 1352
DB 247 ATATATAAATATATTTTATATATTTTAAATATATATAAATATATATTTTATATA 306

QY 1353 ATTACTTTTACACTTTATGTGAAGTGGGTAGGCTGAGTATTTTAAATTTAAATTTAAA 1412
DB 307 TTTATATTTAAATATATATAAATATATTTTATATATTTTAAATATATATAAAT 366

QY 1413 ATTTTAAATTAAGACTATACATAAATTTATGTTTAAAGTTACATTTAAATGAATGCATC 1472
DB 367 ATATTTTATATATTTTAAATATATATAAATATATTTTATATATTTTATATTTTAA 426

QY 1473 ATAACTTTGCCAACATAAACAATAGATAGATACATATGACTTATGAATCGGAGATCA 1532
DB 427 AAATATATAAATATATATTTTATATTTTAAATATATTTGATAAATTTTGATGATATAACCTAA 486

QY 1533 TTTAGTGTGGCCTTTCTTAAAGATTTCAGTTCTGTAAGTAGTCCAGAACTCTCAGTGCC 1589
DB 487 TATAATGTAATAAATTTTAAATATATTTTGAATTTTATTTTGTCTCTTCTATAGTGCC 543

RESULT 13
US-09-949-016-146135
; Sequence 146135, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307

Qy	1233	ATTTTGACAGATGGCTATCAAGACTTGGTAACTATATGATCAAGTGTGTTTTACATAGGGT	1292
Db	187	ATATTTTATATATTTATATTTAAATATATATAAAATATATTTTATATATTTTATATATTTTAAA	246
Qy	1293	AACATTATATATTTAAATTTCCAAAGTGAAGACTTTACAAATGGTGTGATATATATTTTTCCTCC	1352
Db	247	ATATATATAAATATATTTTATATATTTTATATATTTTAAATATATATAAAATATATTTTATATATA	306
Qy	1353	ATTACTTTTTAGACTTTATGTGAAGTGGGTAGGCTGAGTATTTTTTAAATTTTAAAAA	1412
Db	307	TTTATATTTTAAATATATATAAAATATATTTTATATATATTTTAAATATATATAAT	366
Qy	1413	ATTTTAAATTAGAAGCTATCTAAATATGTGTTTAAAGTTACATTTAAATTAATGAATGCATC	1472
Db	367	ATATTTTATATATTTTATATTTTAAATATATATAAAATATATTTTTTATATATTTTATATATTTA	426
Qy	1473	ATAACTTTGCGCAACAATAACACTATAGAGTAGATACATATGACTTTATGAACTGGAGATCA	1532
Db	427	AAATATATATAATATATTTTATATTTTAAATATATTTGATAAATTTGATGATATAACCTAA	486
Qy	1533	TTTAGTGGGCCCTTCTTAAGATTTCCAGTTGTGATAGATAGTGCAGAAATCTCAGTGC	1589
Db	487	TATAATGATATAAATTTTAAATATATTTTGATTTATTTTGTCTTTTGTCTTCTATAGTCC	543

```

US-09-949-016-15851/c
; Sequence 15851, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THERE
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15851
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(205044)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15851

Query Match 1.7%; Score 53; DB 4; Length 205044;
Best Local Similarity 46.8%; Pred. No. 0.063;
Matches 167; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 1233 ATTTTGCAGATGGCTATCAAGACTTGGTAACTATGATCATGTTGTTTTCATAGGGT 1292
Db 200777 ATATTTTATATATTTTATATTTTAAATATATATAAATATATATTTTATATATTTTATA 200777

QY 1293 AACNTATATATTTTATTTTCCAGGTGAAGAACTTACAAATGGTGTGATATATTTTCTCTCC 1352
Db 200717 ATATATATAAATATATTTTATATATTTTATATTTTAAATATATATAAATATATTTTATA 200671

QY 1353 ATTACTTTTGTAGACTTTATCTGAAGGTGGGTAGGCTGAGTATTTTAAATTTTAAAAAAA 1412
Db 200657 TTTATATTTTAAATATATATAAATATATATTTTATATATTTTAAATATATATAAAT 200505

QY 1413 ATTTTAAATTTAGAGCTACTACAAATTTATGTTTAAAGTTTACATTTTAAATTTGATATC 1472

```

Db 200597 ATATTTTATATATTTATATTTTAAATATATATAAATATATATTTTATATATATTTATATTTA 200538
Qy 1473 ATAACTTTGCCCAACAATAACACTATAGAGTAGATACATATGACTTATGAACCTGGAGATCA 1532
Db 200537 AAATATATATAAATATATTTTATATTTTAAATATATTTGATATAAAATTTGATGATATAACCTAA 200478
Qy 1533 TTTAGTGTGGCCTTTCTTAAGATTTTCAGTTTGTAGATAGTCCAGAACTCTCAGTGCC 1589
Db 200477 TATAATGTATAATAAATTTTAAATATATTTTIGATTTATTTTTCCTCTTCTATAGTGCC 200421

Search completed: August 30, 2005, 18:39:39
Job time : 508 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 30, 2005, 07:29:01 ; Search time 4325 Seconds
(without alignments)
4190.116 Million cell updates/sec

Title: US-09-983-000A-2

Perfect score: 1980

Sequence: 1 MRILKRFACIQLLCVCRLD.....GSDQTKHFLTDGYQDLVTI 374

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xl
-Q=/cgn2_1/USPTO_spool_p/US09983000/runat_29082005_155924_14669/app_query.fasta_1.519
-DB=GenEmbl -QFWT=fasap -SUPFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09983000 @CGN 1.1.5600 @runat_29082005_155924_14669 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sw.*

13: gb.un.*

14: gb.vt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1968	99.4	6807	6	CQ725791 Sequence
2	1968	99.4	6924	6	I35776 Sequence 2
3	1968	99.4	7941	6	CQ776705 Sequence
4	1968	99.4	7941	6	AR231695 Sequence

5	1968	99.4	7941	6	AX334330	AX334330 Sequence
6	1968	99.4	7941	9	HUMPTPRZ	M93426 Human prote
7	1968	99.4	7941	11	G20044	G20044 SHSS1987 Er
8	1959	98.9	3350	9	HSU8967	U8967 Human tyros
9	1857	93.8	6801	6	AX683146	AX683146 Sequence
10	1857	93.8	6801	10	RNU04998	U04998 Rattus norv
11	1857	93.8	7851	10	RNU09357	Q09357 Rattus norv
12	1831	92.5	6887	6	CQ777575	CQ777575 Sequence
13	1831	92.5	6887	10	MMU133130	AU133130 Mus muscu
14	1817	91.8	3784	10	MMU428208	AJ428208 Mus muscu
15	1404	70.9	5281	5	BC070850	BC070850 Xenopus 1
16	1376	69.5	5284	5	AB045238	AB045238 Xenopus 1
17	1376	69.5	7725	5	AB045237	AB045237 Xenopus 1
18	1159	58.5	1650	5	BC078205	BC078205 Danio rer
19	649.5	32.8	5403	5	CGU38349	U38349 Gallus gall
20	648.5	32.8	4338	6	AR069915	AR069915 Sequence
21	648.5	32.8	4338	6	BD145125	BD145125 Novel rec
22	648.5	32.8	4338	6	BD145125	I23305 Sequence 1
23	648.5	32.8	4707	9	HUMPTPRG	L09247 Human recep
24	638.5	32.2	2581	10	AY177706	AY177706 Rattus no
25	638.5	32.2	3721	10	AY177705	AY177705 Rattus no
26	638.5	32.2	4391	10	AY177704	AY177704 Rattus no
27	638.5	32.2	4478	10	AY177703	AY177703 Rattus no
28	638.5	32.2	4620	10	MUSPTPRG	L09562 Mouse recep
29	506	25.6	70276	9	AC073095	AC073095 Homo sapi
30	506	25.6	159994	2	AC073471	AC073471 Homo sapi
31	498	25.2	22597	2	AC094129	AC094129 Rattus no
32	482	24.3	5617	10	BC079595	BC079595 Mus muscu
33	479.5	24.2	185859	2	AC133599	AC133599 Mus muscu
34	479	24.2	960	10	S57181	S57181 Ptpg=recp
35	416	21.0	3629	10	BC033432	BC033432 Mus muscu
36	416	21.0	3642	10	BC031385	BC031385 Mus muscu
37	416	21.0	3674	10	BC035941	BC035941 Mus muscu
38	412	20.8	813	6	AR014378	AR014378 Sequence
39	412	20.8	813	6	I33000	I33000 Sequence 9
40	412	20.8	822	6	AR014372	AR014372 Sequence
41	412	20.8	986	6	AR014379	AR014379 Sequence
42	412	20.8	986	6	AR014381	AR014381 Sequence
43	412	20.8	986	6	I32996	I32996 Sequence 3
44	412	20.8	986	6	I33002	I33002 Sequence 12
45	412	20.8	1104	6	AR014371	AR014371 Sequence

ALIGNMENTS

RESULT 1	CQ725791	Sequence	11725 from Patent WO02068579.	6807 bp	DNA	linear	PAT 03-FEB-2004
LOCUS	CQ725791	Sequence	11725 from Patent WO02068579.				
DEFINITION	CQ725791	Sequence	11725 from Patent WO02068579.				
ACCESSION	CQ725791	Sequence	11725 from Patent WO02068579.				
VERSION	CQ725791.1	GI:42287120					
KEYWORDS							
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
REFERENCE							
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.						
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof						
JOURNAL	Patent: WO 02068579-A 11725 06-SEP-2002;						
FEATURES	PF Corporation (NY) (US)						
source	Location/Qualifiers						
	1. .6807						
	/organism="Homo sapiens"						
	/mol_type="unassigned DNA"						
	/db_xref="taxon:9606"						
ORIGIN							
Alignment Scores:							
Pred. No.:	2.88e-178	Length:	6807				
Score:	1968.00	Matches:	372				

```
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 6 Gaps: 0

US-09-983-000A-2 (1-374) x CQ725791 (1-6807)

Qy 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
Db 1 ATGCGAATCCTAAAGCGTTTCTCGCTTGCATTCAGCTCCTCTGTGTTTGCCTGGAT 60

Qy 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40
Db 61 TGGGCTAATGATACTACAGACACAGAGAAACCTTGTTGAAGAGATGGCTGTCTCTAT 120

Qy 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysTyrProThrCysAsnSerProLys 60
Db 121 ACAGGAGCAGTGAATCAAAAAATTTGGGAAAGAAATATCCACATGTAAATAGCCAAA 180

Qy 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
Db 181 CAATCTCTATCAATATTTGATGAAGATCTTTACACAGTAAATGTGAATCTTTAAGAACTT 240

Qy 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db 241 AAATTTTCAGGGTTGGGATAAACAATCATTTGGAAACAACATTCATTCATAACACTGGGAAA 300

Qy 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyValSerGluMetVal 120
Db 301 ACAGTGGAAATTAATCTCAATGACTACCGTGTGCAGCGGAGGATTTTCAGAAATGGTG 360

Qy 121 PheLysAlaSerLysIleThrPheHisTyrTrpGlyLysCysAsnMetSerSerAspGlySer 140
Db 361 TTTAAAGCAAGCAAGATAAATCTTTCAGTGGGAAATGCAATATGTCTCATCTGATGGATCA 420

Qy 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 421 GAGCAGTATTTAGAGGACAAAAATTTCCATTTGAGATGCAAAATCTACTGCTTTGATGCG 480

Qy 161 AspArgPheSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 481 GACCGATTTTCAAGTTTGGAGAGCAGTCAAGAGAAAGGAAAGTAAAGAGCTTTATATCC 540

Qy 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleAspGlyVal 200
Db 541 ATTTGTTTGGAGTTGGGACAGAGAAATTTGGATTTCAAAGCGATTTATTTGATGGATC 600

Qy 201 GluSerValSerArgPheGlyLysGlnAlaAlaLeuAspProPheIleLeuLeuAsnLeu 220
Db 601 GAAAGTGTAGTCGTTTGGAGAGCAGCTGCTTTAGATCCATTCATCACTGTTGACCTT 660

Qy 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProCys 240
Db 661 CTGCCAAACTCAACTGACAGATGATTTACATTTACAATGGCTCATTTGACATCTCTCCCTGC 720

Qy 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 721 ACAGACACAGTGTGATGGATGTTTTTAAAGATACAGTTAGCATCTCTCTGAAAGCCAGTTG 780

Qy 261 AlaValPheCysGluValIleThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 781 GCTGTTTTTGTGAAAGTTCTTCAATATGCAACAATCTGGTTATGTCATGCTGATGACTAC 840

Qy 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 841 TTACAAAAAATTTTCGAGAGCAACAGTACAGTTCTCTAGCAGGTGCTTTTCTCTCATAC 900

Qy 301 ThrGlyLysGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 901 ACTGGAAAGGAGAGATTCATGAGGACGTTTGTAGTTCCAGAACCCAGAAATGTTTCAGGCT 960

Qy 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
Db 941
```

```

Db      481  GACCCGATTTTCAAGTTTTCAGGAGCAGTCAAAGGAAAGGAGGTAAGAGCTTTATCC 540
Qy      181  IleLeuPheGluValGlyThrGluGluLeuLeuAspPheLeuAlaIleLeuAspGlyVal 200
Db      541  ATTTTGTTCAGTTGGGACAGAGAAATTTGGATTTCAAAGCGATATTATGAGGTC 600
Qy      201  GluSerValSerArgPheGlyGlyGlnAlaLeuAspPheIleLeuLeuAsnLeu 220
Db      601  GAAAGTGTTCAGTTGGGAGCAGGCTGCTTTAGATCCATTATCATCTGTTGAACCTT 660
Qy      221  LeuProAsnSerThrAspLysTyrTrpIleTyrAsnGlySerLeuThrSerProCys 240
Db      661  CTGCCAAATCACTCAAGATTAATCAATTTCAATGGCTCAATGACATCTCTCCCTGC 720
Qy      241  ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db      721  ACAGACACAGTTGACTGGATTTGTTTAAAGATACAGTTAGCATCTCTGAAGCCAGTTC 780
Qy      261  AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db      781  GCTGTTTTTGTGAAGTTCTTCAATGCAACCAATCTGGTTATGTCATGCTGATGACTAC 840
Qy      281  LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db      841  TTACAAAACAATTTTCGAGACACAGTACAGTTCTCTAGACAGGTGTTTTCTCTCATAC 900
Qy      301  ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db      901  ACTGGAAGGAAGAGATTCATGACAGTTTGTAGTTCAGAACCAAGAAATGTTCAAGCT 960
Qy      321  AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
Db      961  GACCCAGAGAAATTATACAGCTCTCTGTTACATGGGAAAGACCTCGAGTCGTTATGAT 1020
Qy      341  ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db      1021  ACCATGATTCGAAGTTTGCAGTTTGTACACAGCAGTTGGATGGAGAGACCAACCAAG 1080
Qy      361  HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db      1081  CATGAATTTTTCACAGATGGCTATCAAGACTTGGGTGCTATT 1122

RESULT 3
LOCUS      CQ776705              7941 bp    DNA          linear    PAT 11-MAR-2004
DEFINITION Sequence 391 from Patent EP1394274.
ACCESSION  CQ776705
VERSION    CQ776705.1  GI:45380095
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1. Sugita, Y., Yamaya, M., Kubo, H., Nagai, H. and Izuwara, K.
AUTHORS   Methods of testing for bronchial asthma or chronic obstructive
TITLE     pulmonary disease
JOURNAL   Patent: EP 1394274-A 391 03-MAR-2004;
          Genox Research, Inc. (JP)
FEATURES   Location/Qualifiers
            source
              1..7941
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      3.53e-178      Length:      7941
Score:          1968.00      Matches:    372
Percent Similarity: 99.47%      Conservative: 0
Best Local Similarity: 99.47%      Mismatches: 2
Query Match:    99.39%      Indels:     0

```

```

DB:      6              Gaps:      0
US-09-983-000A-2 (1-374) x CQ776705 (1-7941)
Qy      1  MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
Db      148  ATGCGAATTCCTAAAGCGTTTCTCGCTTGCATTCAGCTCCTCTGTGTTTGGCGCTGGAT 207
Qy      21  TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
Db      208  TGGGCTAATGATATCTACAGCAACAGAGAAACCTTGTGTGAAGAGATTGGCTGTCTCTAT 267
Qy      41  ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
Db      268  ACAGGAGCAGCTGAATCAAAAAAATTTGGGAAAGAAATATCAACATGTAATAGCCCAAAA 327
Qy      61  GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
Db      328  CAATCTCCTATCAATATTATGATGAAGATCTTACACAAGTAATAATGTGAATCTTTAAGAA 387
Qy      81  LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db      388  AATTTTCAGGCTTGGGATTAACAACATCATTTGGAACAACATTCATTCATAACACTGGGAA 447
Qy      101  ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
Db      448  ACAGTGGAAAATTAATCTCCTCAATGACTACCTGTCAGCGGAGGAGTTTCAGAAATGGTG 507
Qy      121  PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
Db      508  TTTAAAGCAAGCAAGATAACTTTTCTACTGGGGAAATGCAATATGTCTCATCTGATGGATCA 567
Qy      141  GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db      568  GAGCATAGTTTAGAAGGACAAAATTTCCACTTGAGATGCAAAATCTACTCTTTGATGGC 627
Qy      161  AspArgPheSerSerPheGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db      628  GACCGATTTTCAAGTTTTCAGGAGAGCAGTCAAAGGAAAGGAGTTTAAGAGCTTTATCC 687
Qy      181  IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleLeuAspGlyVal 200
Db      688  ATTTTGTTCAGGTTGGGACAGAGAAATTTGGATTTCAAAGCGATTTATGATGGAGTC 747
Qy      201  GluSerValSerArgPheGlyLysGlnAlaAlaLeuAspProPheIleLeuLeuAsnLeu 220
Db      748  GAAAGTGTTCAGTTCGTTTGGGAGAGAGCTCTTTAGATCCATTCATCTACTGTTGAACCTT 807
Qy      221  LeuProAsnSerThrAspLysTyrTrpIleTyrAsnGlySerLeuThrSerProCys 240
Db      808  CTGCCAAATCACTCAAGTATTAATCAATTTACAATGCTCATTTGACATCTCTCCCTGC 867
Qy      241  ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db      868  ACAGACACAGTTCAGTTCGATTTGTTTAAAGATACAGTTAGCATCTCTCTGAAAGCCAGTTC 927
Qy      261  AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db      928  GCTGTTTTTGTGAAGTTCTTACATGCAACAAATCTGGTTATGATGCTGCTGATGACTAC 987
Qy      281  LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db      988  TTACAAAACAATTTTCGAGACCAACAGTACAAAGTTCTCTAGACAGGTGTTTTCTCTCATAC 1047
Qy      301  ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db      1048  ACTGGAAGGAAGAGATTCATGAAGCAGTTTGTAGTTTCAAGAACCCAGAAAAATGTTCA 1107
Qy      321  AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
Db      1108  GACCCAGAGAAATTAATACAGCTCTCTTGTATACATGGGAAAGACCTTCGAGTCTGTTATGAT 1167
Qy      341  ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360

```

```
Db 1168 ACCATGATTGAGAGTTTGCAGTTTGTACCGCAGTTGGATGGAGGACCAACCCAG 1227
Qy 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db 1228 CATGAATTTTTCACAGATGGCTATCAAGACITGGGTGCTATT 1269

RESULT 4
AR231695
LOCUS AR231695 7941 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1 from patent US 6455026.
ACCESSION AR231695
VERSION AR231695.1 GI:27272853
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7941)
AUTHORS Mueller, S., Melcher, T. and Chin, D. J.
TITLE Use of protein tyrosine phosphatase zeta as a biomolecular target
in the treatment and visualization of brain tumors
JOURNAL Patent: US 6455026-A 1 24-SEP-2002;
FEATURES Location/Qualifiers
source 1..7941
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3,53e-178 Length: 7941
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservatives: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 6 Gaps: 0

US-09-983-000A-2 (1-374) x AR231695 (1-7941)

Qy 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
Db 148 ATGCGAATCCTAAAGCGTTTCTCGCTTGCATTCAGCTCCTCTGTGTGTTGGCGCTGGAT 207
Qy 21 TrpAlaLeuGlyTyrTyrArgGlnGlnArgLysLeuValGluIleGlyTyrSerTyr 40
Db 208 TGGCTAATGATATCTACAGACAACAGAGAAACTTGTGTGAAGAGATTGGCTGCTCTAT 267
Qy 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
Db 268 ACAGGAGCAGTGAATCAAAAATTTGGGAAAGAAATATCCACATGTAAATAGCCCAAA 327
Qy 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
Db 328 CAATCTCTATCAATATTGATGAAGATCTTACACAAGTAAATGTGAATCTTAAAGAACTT 387
Qy 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db 388 AATTTTCAGGGTGGGATPAAACATCATTTGGAACAACATTCATTCATAACACTGGGAAA 447
Qy 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyValSerGluMetVal 120
Db 448 ACAGTGGAAATTAATCTCACTAATGACTTACCGTGTGCAGCGAGGATTTTCAGAAATGGTG 507
Qy 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
Db 508 TTTAAGCAAGCAAGATAACTTTTTCAGTGGGGAATATGCAATATGTCTCATCTGATGGATCA 567
Qy 141 GluHisSerLeuGluGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 568 GAGCATGTTTAGAAGGCAAAAATTTCCATTTGAGATGCAAACTACTGCTTTTATGCGC 627
Qy 161 AspArgPheSerSerPheGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 628 GACCGATTTTCAAGTTTGTGAGAGCAGTCAAGAGGAAAGGGAAGTTTAAGAGCTTTATCC 687
```

```
Qy 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleLeuAspGlyVal 200
Db 688 ATTTTGTGTTGAGGTTGGGACAGAAAAATTTGGATTTCAAAGCGATTATTGATGGAGTC 747
Qy 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
Db 748 GAAAGTGTAGTCTGTTTGGAGAGAGGCTGCTTTAGATCCATTCACTACATGTTGACCTT 807
Qy 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProCys 240
Db 808 CTGCCAAACTCAACTGACAAGTATTACATTTACAATGGCTCATTTGACATCTCTCCCTCCCTGC 867
Qy 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 868 ACAGACACAGTTGACTGCTGTTGTTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
Qy 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 928 GCTGTTTTTGTGAGGTTCTTACATGCAACATCTGTTATGTATGTCATGCTGATGACTAC 987
Qy 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrIlyPheSerArgGlnValPheSerSerTyr 300
Db 988 TTACAAAACAATTTTCGAGAGCAACAGTACAAGTTCTCTAGACAGGTGTTTTTCTCTCATAC 1047
Qy 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 1048 ACTGGAAGGAAGAGATTCATGAAGCAGTTGTAGTTTCAGAACCAAGAAATGTTTCAGGCT 1107
Qy 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
Db 1108 GACCCAGAGATATATACAGCCTTCTTGTACATGGGAAAGACCTCGAGTCGTTTATGAT 1167
Qy 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 1168 ACCATGATTGAGAAGTTTGCAAGTTTGTACAGCAGTTGGATGGAGAGGACCAACCAAG 1227
Qy 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db 1228 CATGAATTTTTCAGATGGCTATCAAGACTTTGGGTGCTATT 1269

RESULT 5
AR2334330
LOCUS AR2334330 7941 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 4839 from Patent WO0194629.
ACCESSION AR2334330
VERSION AR2334330.1 GI:18125049
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 4839 13-DEC-2001;
FEATURES Avalon Pharmaceuticals (US)
source Location/Qualifiers
1..7941
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 3,53e-178 Length: 7941
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservatives: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 6 Gaps: 0
```

US-09-983-000A-2 (1-374) x AX3334330 (1-7941)	
QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20	1168 ACCATGATTGAGAGCTTTTGCAGTTTGTACAGCAGTTGGATGGAGACCAACCAAG 1227
DB 148 ATGCGAATCTAAAGGTTTCTCGCTTGCAATTCAGCTCTCTGTGTTTGCAGCTCGAT 207	361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
QY 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluIleGlyTrpSerTyr 40	1228 CATGAATTTTTCACAGATGGCTATACAGACTTGGGTGCTATT 1269
DB 208 TGGGCTAAATGGATCTACACACAAGAGAAATCTGTTGAAGAGATTGGCTGCTCTAT 267	RESULT 6
QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLeuTyrProThrCysAsnSerProLys 60	HUMPTPRZ 7941 bp mRNA linear PRI 08-JAN-1995
DB 268 ACAGGAGCACTGAATCAAAAAATTCGGGAAGAAATATCCAACATGTATAGCCCCAAA 327	LOCUS HUMPTPRZ
QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysValLeu 80	DEFINITION Human protein tyrosine phosphatase zeta-polypeptide (PTPRZ) mRNA,
DB 328 CAATCTCTCATATATGATGAGATCTTACCAAGTAAGTAATCTTAAGAAACTT 387	complete cds.
QY 81 LysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100	ACCESSION M93426
DB 388 AAATTTTCAGGTTGGGATAAAACATCATTTGGAAAAACACATTCATTCAATCACTGGGAA 447	VERSION M93426.1 GI:190743
QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyValSerGluMetVal 120	KEYWORDS carbonic anhydrase-related transmembrane protein; protein tyrosine phosphatase zeta-polypeptide.
DB 448 ACAGTGGAAATTAATCTCAATGACTACCGTGTGAGCGAGGAGTTTCAGAAATGGTG 507	SOURCE Homo sapiens
QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140	ORGANISM Homo sapiens
DB 508 TTTAAGCAAGCAAGATAACTTTTCTCGGGGAAATGCAATATGTCATCTGATGATCA 567	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160	REFERENCE 1 (bases 1 to 7941)
DB 568 GAGCATAGTTTAGAGGACAAAAATTTCCACTTGAGATGCAATCTACTGCTTTGATGGG 627	AUTHORS Krueger,N.X. and Saito,H.
QY 161 AspArgPheSerSerPheGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180	TITLE A human transmembrane protein-tyrosine-phosphatase, PTP zeta, is expressed in brain and has an N-terminal receptor domain homologous to carbonic anhydrases
DB 628 GACCGATTTTCAAGTTTTCAGGAAGCAGTCAAGGAAAGGAAGTTAAGACTTTATCC 687	JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (16), 7417-7421 (1992)
QY 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleLeuAspGlyVal 200	MEDLINE 92366472
DB 688 AITTTGTTTGAGGTTGGGACAGAGAAATTTGGATTTCAAGCGCATTAATGATGGAGTC 747	COMMENT 1323835
QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220	FEATURES
DB 748 GAAAGTTTAGTGTGTTTCGGAGCAGGCTGCTTTAGATCCATTCACTGTTGACCTT 807	source
QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProCys 240	1. 7941
DB 808 CTGCCAAACTCAACTGACAAGTATTACATTTACAATGGCTCATTCACATCTCTCCCTGC 867	/organism="Homo sapiens"
QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260	/mol_type="mRNA"
DB 868 ACAGACACAGTTGACTGGATTTGTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927	/db_xref="taxon:9606"
QY 261 AlaValPheCysGluValLeuThrMetGlnSerGlyTyrValMetLeuMetAspTyr 280	/map="Unassigned"
DB 928 GCTGTTTTTGTGAGTTCTTACATGCAACAATCTGTTATGTCATGCTGATGACTAC 987	/tissue_type="brain"
QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300	/dev_stage="fetus"
DB 988 TTACAAAACAAATTTTCAGAGCAACAGTACAAGTTCTCTAGACAGTGTGTTTCTCTCATAC 1047	/tissue_lib="lambda gt10 and gt11"
QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320	1. 7941
DB 1048 ACTGAAAGGAGAGATCATCAGACAGTTTGTAGTTCAGAACCAAGAAATGTTCCAGCT 1107	/gene="PTPRZ"
QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluAtgProArgValValTyrAsp 340	148. .7092
DB 1108 GACCCAGAGAAATATACAGGCTTTCTGTTATCATGGGAAAGACCTCGAGTCTGTTATGAT 1167	/gene="PTPRZ"
QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360	/note="carbonic anhydrase-like domain"
	/product="protein tyrosine phosphatase zeta-polypeptide"
	/protein_id="AAA60225.1"
	/db_xref="GI:190744"
	/db_xref="GDB:G00-127-353"
	/translation="MRILKRFPLACIQLLCVCLDWANGYVQRKLVVEEIGMSVTGAL
	NOKNWKYPTCNPKSPKIPINIDSLQVNVNLKAKFGQWDTKTSLENTFTHNTGKTV
	EINLTNDYRVSGGVSEVSEVFKASKITPHGKNMSSDQSEHSLGKQPLENQLVCFDA
	DRFSSPREAVKGGKLRALSILFEVGTENLDKRAIDGVSESVRFQKQALDPFILL
	NLLPNSDTKYYIINGSLTSPPTDVTDMIVPKDTVSISESQAVFCEVLVTMQGGYVM
	LMYDQNNFREQQYKFRQVPSYTSYTGKEIHEAVCSSEPVAVCPENYTSLLVWTR
	PRVYDITWIKFAVLVYQQLDGEDOTHEFLTQYDQILGAILNNKNSVYLVQVIAIC
	TNGLYGYSQDLIVDMPTDNPDLPLFLIGTERIIKEEREGKDIIEGAIVNPGRDA
	TNQRKKEPQISITTHNRICTKNEAKTRSPTRGSEPSGKGVDPNTSLNSTQPVIT
	KLATEKISLTSVTPHVEGTSASLNDGSKTVLRSPHMSLSTGSAESLNTSVIT
	EYEEESLITFKLDTGAESSGSPATSAIPFISENISQGYIFSSENPETITVDLIP
	ESARNAEDSTSGSESLKDPSEGNVWFPSSDTDAQPDVSGRSFLQNTYETIR
	VDESEKTKFSFAGVMSQGSVTDLEMPHYSTFAYPTVPHAFTPSSRQDLVST
	VNVVYQOTPVYNGETPLQPSYSSEVPLVTLPLLDNLQILNTTPAASSSDSLHATP
	VFSPVDSVPSRILSSYDGAELLPPSSASFSSELFRHLHTVSQLIPQVTSATESDKVPL
	HASLPVAGDLLLEPSLAQYDVLSTTHAASELTFGSESGVLTKLTFMFSQVEPSSD
	AMHARSGGPEPSYALSDNGSGSHFTVSYSSAIPVHDSVGVTVQSLFGSPSHPIIP
	KSLITPTASLLOPTHALSDGSGWSSDSBFLPDTDGLTALNLISSPVSAEFTYT
	TSVFGDNDKNNIYGNBTELQIPSENVVYPSSTVMPNNYDVNVLKNASLIQET
	SVSTSSYKGMFPGKSLAHTTKVPDHEISQVPENNVFSQPTHTVYVSAQSDTSLKPLVA
	NSEPADSDPASSEMSPSTQLLPYETSASFTEVLQPSFQASDVDTLLKTVLPAPVS
	DPILVETPKVDKISSTMLHLIVNSASSENMLHSTVVFVDVSPSHHVASLQGLTI
	YSASEKPEVLLKSSSHQVVPVSLYSNDELFTQNLANLEINQAHQPSRHVFATPVLIS
	EPLNTLNLKLIHSDIILTSTKSVTKVFPAGIPTVASDTPFVSTDHSVPIGNGHVAITA
	VSPHRDGSVTSKLLPSPKATSELSHSAKSDAGLVGGGEGDGTDDDDDDDDDRSDGL

SIHKMCSSEYBESQKVMNDSDTHENSMDNDNPNISVLSSENSEEDNRVTSVSSDSQ
 TGMRSPPKSPANGSLQKNDGKRENDIQTSALLSPESKAWALITSDSESGSQ
 GTSLSLNETSTDFSAADTNEKADGILAAAGSEITPGFPQSPSTSVTSSESVFV
 SEASASNSHESRIGLAEGLSEKAVIPLVIVSALTFCILVILVIGLILYWRKCFQTA
 HFYLEDSTSPVPTPIPIISDVVGAIPIHFPHKVAIDLHASSGTFBEFLEKIF
 YQVQSCVTDIGITADSSNHPDNKHKNYINIVADHSRVKLAOLAKDQKGLTDYINA
 NYVDGNPKAYIAAQQPLKSTAEDEFWRMIWEHNEVEVIMTNLVEGRKCDQYWA
 DGSBEYNFLVTQKSQVLAITYTVNFTLRNFKIKGSKQGRPSGRVVTQHYTQWPD
 MGVEYSLPVLTFVRKAAIARHAGVGVVHCSSAGVGTGYIVLSDMLQIOHEGT
 NIFGLKHIRSQRYNLVQTEQYVFIHDTLVEALISKETEVLDHSHIAYNALLIPG
 AGTKLEQFOLLQSQSDYSAALKQCNREKNTSSIIIPVRSRVGSSISGEGT
 DYINASYTMGYQSNERTIOTPHLLHTIKDFWRMIWDHNAOLVVMIDPQGNMADREFV
 WYMPKDEPNCESKPTLMAEHEKLSNEEKLIIODPLEATODDYLVLEVRHFCQPM
 PNPSPIKTELLSVIKEEAANDGPMIVDEHGGVGTAGTFCALTILMQLKENSIV
 DYOQVAKRMNLMPGVFADIEQYQFLYKVLISLVSSTQENPSTSLDSNGAALPDGNI
 AESLESIV"

ORIGIN

Alignment Scores:
 Pred. No.: 3,53e-178 Length: 7941
 Score: 1968.00 Matches: 372
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 2
 Query Match: 99.39% Indels: 0
 DB: 9 Gaps: 0

US-09-983-000A-2 (1-374) x HUMPTPRZ (1-7941)

Qy 1 MetArgilleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
 Db 148 ATGCAATCTTAAGCGTTTCTCGCTTGCAATCAGCTCCTCTGTGTTGCGGCTGGAT 207
 Qy 21 TrpAlaAenGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40
 Db 208 TGGGCTAATGATGATACACAGCAACAGAGAAATCTGTGTAAGAGATGGCTGCTCAT 267
 Qy 41 ThrGlyAlaLeuAenGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProlys 60
 Db 268 ACAGGAGCACTGAATCAAAATAATTTGGGGAAGAAATAATCAACATGTAATAGGCCAAA 327
 Qy 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
 Db 328 CAATCTCTATCAATATTGATGAGATCTTACACAGTAATGTAATGTAATCAAGAACTT 387
 Qy 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
 Db 388 AAAATTCAGGGTTGGGATAAACAATCATTTGGAACAAACATTCATTCATAACACTGGAAA 447
 Qy 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyValSerGluMetVal 120
 Db 448 ACAGTGGAAATTAATCTCCTCAATGACTACCGGTGTCAGCGAGGAGTTTCAGAAATGGTG 507
 Qy 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
 Db 508 TTTAAGCAAGCAAGATTAATCTTCACTGGGAAAATCAATATGTCATCTGATGGATCA 567
 Qy 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
 Db 568 GAGCATAGTTTAGAAGGAGCAAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATGCG 627
 Qy 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
 Db 628 GACCGATTTCAAGTTTGTAGGAAGCAGTCAGAGGAAAGGAAGGAGTTTAAGAGCTTTATCC 687
 Qy 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleAspGlyVal 200
 Db 688 ATTTGTTTGGAGTTGGGACAGAGAAAATTTGGATTTCAAGCGATTAATTGATGGATC 747
 Qy 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
 Db 748 GAAAGTGTAGTCGTTTGGGAAGCAGGCTGCTTTAGATTCATATTCATATCTGTTGAACCTT 807
 Qy 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240

Db 808 CTGCCAAATCACTGACAAAGTATTACATTTACATGGCTCATTCATCTCCTCCCTGC 867
 Qy 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
 Db 868 ACAGACACAGTTGACTGGATTGTTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
 Qy 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
 Db 928 GCTGTTTTTGTGAAGTCTTTACAATGCAACAATCTGGTTATGTCATGCTGATGGACTAC 987
 Qy 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
 Db 988 TTACAAAACAATTTTCAGAGCAACAGTACAGTTCTCTAGACAGGTGTTTTCTCATAC 1047
 Qy 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
 Db 1048 ACTGGAAGGAAGAGATTTCATGAAGCAGTTTGTAGTTTCAGAACCAAGAAATGTTCCAGCT 1107
 Qy 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
 Db 1108 GACCAGAGAATATATACCAGCCTTCTTTGTACATGGGAAAGACCTCGAGTCTGTTATGAT 1167
 Qy 341 ThrMetIleGlyLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
 Db 1168 ACCATGATTGAGAAGCTTTTGACAGTTTGTACAGCAGTTGGATGGAGAGGACCAACCAAG 1227
 Qy 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
 Db 1228 CATGAATTTTGCAGATGCTATCAAGACTTGGGTGCTATT 1269

RESULT 7

G20044 LOCUS 7941 bp DNA linear STS 28-SEP-1998
 DEFINITION SWSS1987 Eric D. Green Homo sapiens STS genomic, sequence tagged site.

ACCESSION G20044
 VERSION G20044.1 GI:1254743
 KEYWORDS STS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 7941)
 Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F., Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulcon,R.S., Leckie,M.P. and Green,E.D.
 A collection of 1814 human chromosome 7-specific STSs
 Genome Res. 7 (1), 59-64 (1997)
 97189344
 9037602
 2 (bases 1 to 7941)
 Green,E.D.
 Human chromosome 7 STSs (1997)
 Unpublished (1997)
 Synonyms: PTPRZ
 GDB: GDB:3754378
 GDB_PSEG: PTPRZ
 Contact: Eric D. Green
 Genome Technology Branch
 National Human Genome Research Institute/NIH
 49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
 Tel: 3014020201
 Fax: 3014024735
 Email: egreen@nhgri.nih.gov
 Primer A: CTGGTCTTACTCTACCAAG
 Primer B: GAACAATGTTTGTAAGGTG
 STS size: 152
 PCR Profile:
 Presoak: 0 degrees C for 0.00 minute(s)
 Denaturation: 92 degrees C for 0.17 minute(s)
 Annealing: 55 degrees C for 1.00 minute(s)
 Polymerization: 72 degrees C for 1.00 minute(s)

PCR Cycles: 35
Thermal Cycler: PerkinElmer 9600

Protocol:
Template: 30-100 ng
each 1 uM
Primer: each 200 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer:
MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

This STS was developed from sequence determined by another investigator. See GenBank record: M93426 for additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgri.nih.gov/DIR/STB/CHR7>. Also see Genomics 11:548-64 (1991) [MUID=92128937].

FEATURES

source
1..7941
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="7"
/clone_lib="Eric D. Green"
1..7941
/gene="PTPRZ"
7705..7856
/gene="PTPRZ"
7705..7722
/gene="PTPRZ"
complement(7837..7856)

Alignment Scores:

Pred. No.: 3.53e-178 Length: 7941
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 11 Gaps: 0

US-09-983-000A-2 (1-374) x G20044 (1-7941)

Qy 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
Db 148 ATGCGAATCCTAAAGCGTTCTCGCTTGCAATTCAGCTCCTCTGTGTGTGGCGCGCTGGAT 207
Qy 21 TtpAlaAenGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40
Db 208 TGGCGCTAATGGGATCTACAGACACAGAGAAACTGTGTGAAGAGATGGCTGGTCCCTAT 267
Qy 41 ThrGlyAlaLeuAenGlnLysAenTyrGlyLysLysTyrProThrCysAenSerProLys 60
Db 268 ACAGGAGCACTGAATCAAAAATTTGGGAAAGAAATATCCACATGTAATAGCCAAAA 327
Qy 61 GlnSerProIleAenIleAenGluAspLeuThrGlnValAenValAenLysLysLeu 80
Db 328 CAATCTCCTATCAATATTTGATGAGATCTTACACAAGTAAATGTGAATCTTAAGAAACTT 387
Qy 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAenThrPheIleHisAenThrGlyLys 100
Db 388 AAATTTTCAGGGTGGGATAAACATCATTTGGAACAACATTCATTCATAACACTGGGAAA 447
Qy 101 ThrValGluIleAenLeuThrAenAspTyrArgValSerGlyValSerGluMetVal 120
Db 448 ACAGTGGAAATTAATCTCAATGACTACCGTGTGACGGAGGATTTCAAGAAATGGT 507
Qy 121 PheLysAlaSerLysIleThrPheHisTyrGlyLysCysAenMetSerSerAspGlySer 140
Db 508 TTTAAAGCAAGCAAGATAACTTTTCACTGGGGAAATGCAATATGCTATCTGATGATCA 567

Qy 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 568 GAGCATAGTTTAGAAGGACAAAATTTCCACTTGAGATGCAATCTACTGCTTTGATGCG 627
Qy 161 AspArgPheSerSerPheGluGluAlaValysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 628 GACCGATTTTCAAGTTTTGAGGAAGCAGTCAAGGAAAAGGAAGTTAAGAGCTTTATCC 687
Qy 181 IleLeuPheGluValGlyThrGluGluAenLeuAspPheLysAlaIleAspGlyVal 200
Db 688 ATTTTGTGTGAGGTGGGACAGAAAGAAATTTGGAATTTCAAAGCGAATTAATGATGAGTC 747
Qy 201 GluSerValSerArgPheGlyLysGlnAlaAlaLeuAspPropheIleLeuLeuAenLeu 220
Db 748 GAAAGTGTAGTCTGTTTTGGGAGCAGGCTGCTTTAGATCCATCTACTACTGTTGACCTT 807
Qy 221 LeuProAenSerThrAspLysTyrTyrIleTyrAenGlySerLeuThrSerProCys 240
Db 808 CTGCCAAACTCACTGACCAAGTATTACATTTACAAATGGCTCATTTGACATCTCTCCCTGC 867
Qy 241 ThrAspThrValAspTyrIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 868 ACAGACACAGTTGACTGGATTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
Qy 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 928 GCTGTTTTTGTGAAGTTCTTACAAATGCAACATCTGTTATGTCATGCTGATGACTAC 987
Qy 281 LeuGlnAenAenPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 988 TTACAAAACAATTTTCGAGAGCAACAGTACAAGTTCTCTAGACAGGTGTTTTCTCTCATC 1047
Qy 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAenValGlnAla 320
Db 1048 ACTGAAAGGAAGAGATTCATGAAGCAGTTGTAGTTTCAAGAACAGAAAATGTTTCAGCT 1107
Qy 321 AspProGluAenTyrThrSerLeuLeuValThrTyrGluArgProArgValValTyrAsp 340
Db 1108 GACCCAGAGAAATATACAGCCTTCTGTTTACATGGGAAAGACCTCGATGCTTTATGAT 1167
Qy 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 1168 ACCATGATTGAGAAAGTTTGCAAGTTTTGTACCAGCAGTTGGATGGAGAGGACCAACCAAG 1227
Qy 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db 1228 CATGAATTTTTCAGACATGGCTATCAAGACTTGGGTGCTATT 1269

RESULT 8

HSU88967 3350 bp mRNA linear PRI 08-AUG-1998
LOCUS Human tyrosine phosphatase zeta polypeptide 2 (HTPZP2) mRNA,
partial cds.
DEFINITION U88967
ACCESSION U88967
VERSION U88967.1 GI:3282240
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM
REFERENCE Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 3350)
Ouyang, P., Lubyova, B., Gardellin, P., Kurzbauer, R. and Weich, A.
TITLE Molecular cloning and expression analysis of five novel genes in
chromosome 19p36
JOURNAL Genomics 50 (2), 187-198 (1998)
MEDLINE 96317532
PUBMED 9653645
REFERENCE 2 (bases 1 to 3350)
Ouyang, P., Lubyova, B., Gardellin, P., Kurzbauer, R. and Weich, A.
TITLE Direct Submission
AUTHORS Submitted (10-FEB-1997) Research Institute of Molecular Pathology,
Dr. Bohr Gasse 7, Vienna A-1030, Austria
JOURNAL

Qy	141	GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla	160
Db	756	GAGCATAGTTTAGAAGGACAAAATTTCCACTTGGAGATGCAAAATCTACTGCTTTGATGG	815
Qy	161	AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer	180
Db	816	GACCGATTTTCAAGTTTGTAGGAAGCAGTCAAAAGGAAAAAGGGAAGTTTAAAGAGCTTTATCC	875
Qy	181	IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIleAspGlyVal	200
Db	876	ATTTTGTGTTGAGGTTGGGACAGAGAAAATTTGGATTTCAAAGCGATTATTGATGAGATC	935
Qy	201	GluSerValSerArgPheGlyLysGlnAlaAlaLeuAspProPheIleLeuLeuAsnLeu	220
Db	936	GAAGTGTGTAGTTCGTTTGGGAAGCAGCGCTGCTTTAGATCCATTTCATCTGTTGAACCTT	995
Qy	221	LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys	240
Db	996	CTGCCAAACTCACTGCAAGATATTACATTTACAATGGCTCATTTGACATCTCTCCCTCGTCG	1055
Qy	241	ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu	260
Db	1056	ACAGACACAGTTGACTGGATTGTTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG	1115
Qy	261	AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr	280
Db	1116	GCTGTTTTTGTGAAGTTCCTTACAAAGCAACAATCTGTTATGTCATGCTGATGGACTAC	1175
Qy	281	LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr	300
Db	1176	TTACAAAACAATTTTCGAGAGCAACAGTACAGATTCTCTAGACAGGTGTTTTCCTCATAC	1235
Qy	301	ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla	320
Db	1236	ACTGGAAGGAAGAGATTTCATGAAGCAGCTTGCAGGTGAGAACACAGAAAATGTTTCAGGCT	1295
Qy	321	AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp	340
Db	1296	GACCCAGAGAAATATATACAGCCCTTCITTTGTACATGGGAAAGACCTCGAGTCGTTTATGAT	1355
Qy	341	ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys	360
Db	1356	ACCATGATTGAGAAGTTTGCAGTTTTGTACCAGCAGTTGGATGGAGAGGACCAACCAAG	1415
Qy	361	HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle	374
Db	1416	CATGAATTTTTCAGACAGATGGCTATCAAGACTTGGGTGCTATT	1457
RESULT 9			
LOCUS	AX683146	6801 bp DNA linear	PAT 29-MAR-2003
DEFINITION	Sequence 120 from Patent EP1279744.		
ACCESSION	AX683146		
VERSION	AX683146.1 GI:29370167		
KEYWORDS	Rattus norvegicus (Norway rat)		
SOURCE	Rattus norvegicus		
ORGANISM	Rattus norvegicus		
REFERENCE	1. .6801		
AUTHORS	Brookebank, R.A., Dixon, A.K., Lee, K. and Pinnock, R.D.		
TITLE	Identification and use of molecules implicated in pain		
JOURNAL	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
FEATURES	1. .6801		
source	/organism="Rattus norvegicus"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:10116"		
ORIGIN	/note="Phosphacan"		

FEATURES	Location/Qualifiers
source	1..3350
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/chromosome="1"
	/map="1p36"
gene	1..>3350
	/gene="HTP2P2"
CDS	336..>3350
	/gene="HTP2P2"
	/note="similar to protein tyrosine phosphatase"
	/codon_start=1
	/product="tyrosine phosphatase zeta polypeptide 2 HTP2P2"
	/protein_id="AAC39934.1"
	/db_xref="GI:3282241"
	/translation="MRILKRFLLACIQLLCVRLDWANGYYRQRKLVEEIGWSYTGAL NOKWKKYPTCNSPKQSPINIDDLTQVNVNLKLFQGWDTKSLNTFPHNTKTV EINLTNDYRVSGVSEWVFKASKITFHGKNMSSDGSSEHSLEQKFLPMQLYCPDA DRPSPEAVGKGLKRLALSILFVGTENLDPKAIIDGVESVSRFGKQALDPFILL NLLPNSDKYIYISLTSPPCTDVTDMIVFKOTVSISESQLAFCVLTMQQSGYVM LMDYQNNFRQQYKFRQFSSYTGKEEIEHAACRSEPNVQADPNYTSLSLVTWR PRVVDYTMIEFAVLQQLDGEDQTKHEFLTDGYQDLGAILNNLLPNMSYVLQIVATC TNGLYKYSDDLI VDMPTNDPELDFPELIGTERIIKEEBEGKIDIBERGAVNPRDSA TNQIRKKEPOISTTTHNRICTKNEAKTNRSPTRGSEFGKGDVPNTSLNSTSQPT KLATEKILSNTQVTELPHTVGTSASINDGSKTVLRSPHMLSGTAGSLNTYSIT EYESLSLTSFKLDTGSDSGSSPATSAPFISENISQGYIFSSENPETIYDVLPI ESARNAGEDSTSEBSLEKDPMSGNVWPFSSDTIAQPDVGGRESPIQNTYETIR VDESEKTTKSPAGVMSQPSVTDLEMPHYSTPAYFPTVTHAFTPSRQDLVST VNVVYSSOTQVNGEYTPLOPSYSEVFPVLTPLLDNQILNTTTPAASSTDSALHATP VFPSPVDSFESLISYDGAPLLPFSSASFSSELFRHLHTVSQILPQVTSATESDKVPL HASLPVAGGDLILLSPSLAQYSDVLSTTHAASETLEFGSESGVLYKTLMFQVPEPSSD AMHARSGSPFSFALSDNBSQHIFTVSYSSAIPVHDSVGVTTQGSLLTQSLFSGPSHIP KSLITPTASLLQPTHALSGDGSWGAFFDKL"
ORIGIN	
Alignment Scores:	
Pred. No.:	8,296-178
Score:	1959.00
Percent Similarity:	98.93%
Best Local Similarity:	98.93%
Query Match:	98.94%
DB:	9
US-09-983-000A-2 (1-374) x HSU89967 (1-3350)	
Qy	1 MetArgIleLeuLysAtrgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
Db	336 ATGCGAATCCTAAAGCGTTTCTCGCTTGCAATCAGCTCCCTCTGTGTTGCGCGCTGGAT 395
Qy	21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
Db	396 TGGGCTAATGGATACTACAGACAACAGAGAAACATTTGTTGAAGAGATTGGCTGGTCTTAT 455
Qy	41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
Db	456 ACAGAGCACTGAAATCAAAAAAATTGGGGAAGAAATATCAACATGTAATAGCCCAAAA 515
Qy	61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
Db	516 CAATCTCTATCAATATTGATGAGATCTTACACAAGTAATGTAATCTTAAAGAAACTT 575
Qy	81 LysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db	576 AAATTTTCAGGTGGGATAAACAATCATTTGGAAAAACACATTCATTTCATPAACACTGGGAAA 635
Qy	101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyValSerGluMetVal 120
Db	636 ACAGTGGAAATTAATCTCACTAATAGACTACCGTGTACGGAGGAGGATTTTCAAGAAATGGTG 695
Qy	121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
Db	696 TTTTAAAGCAAGCAAGATAACTTTTCACTGGGGAAAAATGCAATATGTCATCTCATGGATCA 755

Alignment Scores:
 Pred. No.: 1,3e-167 Length: 6801
 Score: 1857.00 Matches: 346
 Percent Similarity: 96.26% Conservative: 14
 Best Local Similarity: 92.53% Mismatches: 14
 Query Match: 93.79% Indels: 0
 DB: 6 Gaps: 0

US-09-983-000A-2 (1-374) x AX683146 (1-6801)

QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
 DB 106 ATGCGAATCTCGACAGACTTCCTCGCGTGCCTTACGCTACTGCGGTGTGTGCGCTGGAC 165
 QY 21 TrpAlaAsnGlyTyrTrpArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
 DB 166 TGGCTTATGGATACTACAGACACAGAGAACTGTGTGAAGAGATTGGCTGGTCTCTAT 225
 QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
 DB 226 ACAGGAGCACTAAATCAAAAATTTGGGGAAGAAATATCCAAATATGTAATAGCCCAAG 285
 QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
 DB 286 CAGTCTCTATTAAATTTGATGAAGATCTTACACAAGTAAATGTGAATCTTAAAGAAACTG 345
 QY 81 LysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleAsnThrGlyLys 100
 DB 346 AAATTTCAAGGTGGGAAACCACTCTTGGGAAACAGTTTATTCACAACTGGGAAA 405
 QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyLysValSerGluMetVal 120
 DB 406 ACAGTGGAAATAAATCTCCCAATGACTACTATCTCAGTGGAGGACTTTCAGAAAGGTC 465
 QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
 DB 466 TTCAAGGCAAGCAAAATGACTTTTCACTGGGGGAAATGCAATGTGTCTCGAAGAGATCG 525
 QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
 DB 526 GNAACAGCTTAGAAGGCGGANGTTTCCACTCGAGATGCAAAATCTACTGCTTCGATCGG 585
 QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
 DB 586 CACCGATTTCAAGTTTGGAGAAACAGTTAAAGGAAAGGAGGTAAAGGCTCTATCC 645
 QY 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIleAspGlyVal 200
 DB 646 ATTTTATTTGAGATTGGAGTTTGAAGAAAATTTGGATTACAAAGCCATTTATTGATGGGACT 705
 QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspPropheIleLeuLeuAsnLeu 220
 DB 706 GAAAGTTGTAGTCGTTTGGAAAGCAGGCTGCTTAGATCCGTTTCATCTTGAGAACCTC 765
 QY 221 LeuProAsnSerThrAspLysTyrTrpIleTyrAsnGlySerLeuThrSerProCys 240
 DB 766 CTGCCAACTCCACTGACAGTACTACATTTACACAGGATCACTGACATCCCTCCCTGC 825
 QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
 DB 826 ACAGACACGTCGAATGGATTGTTTAAAGGATACAGTTAGCATCTCTGAAAGCCAGCTG 885
 QY 261 AlaValPheCysGluValLeuThrMetGlnSerGlyTyrValMetLeuMetAspTyr 280
 DB 886 GCTGTATTTTGGAAGTTCTCAAAATGCAACAGTCCGGGTATGTATGTTGATGATTAC 945
 QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
 DB 946 TTACAAAACATTTCCGAGAACACAGTACAAAGTTTCCAGGCGAGGTGTTTCTTCGTAC 1005
 QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320

DB 1006 ACTGGAGGAGGAGATTTCATGAAGCAGTGTGTCAGTTCAGAACACAGAAAATGTGCAAGCT 1065
 QY 321 AspProGluAsnTyrThrSerLeuValThrTrpGluArgProArgValValTyrAsp 340
 DB 1066 GACCTGTGAATTTACACAGGCTTCTGATCATCTGGGAAAGGCTTCGGGTTGTTTATGAC 1125
 QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
 DB 1126 ACAATGATTGAGAAGTTTGCCTTCTGTACCAAGCCTCTGGAGGGAACACCAACCAAG 1185
 QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
 DB 1186 CATGAGTTTTTAACAGATGGCTATCAGGACTTGGGTGCCATT 1227

RESULT 10
 RNU04998 6801 bp mRNA linear ROD 02-FEB-1996
 LOCUS Rattus norvegicus Sprague-Dawley phosphacan mRNA, complete cds.
 DEFINITION U04998
 ACCESSION U04998.1 GI:461371
 VERSION
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Rattus norvegicus
 Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 6801)
 AUTHORS Maurel, P., Rauch, U., Flad, M., Margolis, R.K. and Margolis, R.U.
 TITLE Phosphacan, a chondroitin sulfate proteoglycan of brain that
 interacts with neurons and neural cell-adhesion molecules, is an
 extracellular variant of a receptor-type protein tyrosine
 phosphatase
 J. Biol. Chem. 269 (16), 2512-2516 (1994)
 JOURNAL MEDLINE 94195772
 PUBMED 7511813
 REFERENCE 2 (bases 1 to 6801)
 AUTHORS Grumet, M., Milev, P., Sakurai, T., Karthikeyan, L., Bourdon, M.,
 Margolis, R.K. and Margolis, R.U.
 TITLE Interactions with tenascin and differential effects on cell
 adhesion of neurocan and phosphacan, two major chondroitin sulfate
 proteoglycans of nervous tissue
 J. Biol. Chem. 269 (16), 12142-12146 (1994)
 JOURNAL MEDLINE 94216329
 PUBMED 7512960
 REFERENCE 3 (bases 1 to 6801)
 AUTHORS Engel, M., Maurel, P., Margolis, R.U. and Margolis, R.K.
 TITLE Chondroitin sulfate proteoglycans in the developing central nervous
 system. I. cellular sites of synthesis of neurocan and phosphacan
 J. Comp. Neurol. 366 (1), 34-43 (1996)
 JOURNAL MEDLINE 97020465
 PUBMED 8866844
 REFERENCE 4 (bases 1 to 6801)
 AUTHORS Meyer-Puttlitz, B., Junker, B., Margolis, R.U. and Margolis, R.K.
 TITLE Chondroitin sulfate proteoglycans in the developing central nervous
 system. II. Immunocytochemical localization of neurocan and
 phosphacan
 J. Comp. Neurol. 366 (1), 44-54 (1996)
 JOURNAL MEDLINE 97020466
 PUBMED 8866845
 REFERENCE 5 (bases 1 to 6801)
 AUTHORS Milev, P., Friedlander, D.R., Sakurai, T., Karthikeyan, L., Flad, M.,
 Margolis, R.K., Grumet, M. and Margolis, R.U.
 TITLE Interactions of the chondroitin sulfate proteoglycan phosphacan,
 the extracellular domain of a receptor-type protein tyrosine
 phosphatase, with neurons, glia, and neural cell adhesion molecules
 J. Cell Biol. 127 (6 Pt 1), 1703-1715 (1994)
 JOURNAL MEDLINE 95096181
 PUBMED 7528221
 REFERENCE 6 (bases 1 to 6801)
 AUTHORS Milev, P., Meyer-Puttlitz, B., Margolis, R.K. and Margolis, R.U.
 TITLE Complex-type asparagine-linked oligosaccharides on phosphacan and
 protein-tyrosine phosphatase-zeta/beta mediate their binding to
 neural cell adhesion molecules and tenascin

Db	226	ACAGGAGCCTAAATCAAAAAATTTGGGGAAGAAATATCCAATATGTAAATAGGCCAAAG	285
Qy	61	GlnSerProIleAenIleAaspGluAaspLeuThrGlnValAenValAenLeuIleAenLeu	80
Db	286	CAGTCTCTCTATTAATATTGATGAAGATCTTACACAAGTCAATGTGAATCTTTAAGAAAC	345
Qy	81	LysPheGlnGlyTrpAaspLysThrSerLeuGluAenThrPheIleHisAenThrGlyLys	100
Db	346	AAATTTCAGGGTTGGGAAAAGCGTCTTGGAAAACACGTTTCATTCACAACTGGGAAA	405
Qy	101	ThrValGluIleAenLeuThrAenAaspTrpArgValSerGlyValSerGluMetVal	120
Db	406	ACAGTGGAAATAAATCTCACAATGACTACTATCTCAGTGAGGACCTTTCAGAAAAGGTC	465
Qy	121	PhelysAlaSerLysIleThrPheHisTrpGlyLysCysAenMetSerSerAaspGlySer	140
Db	466	TTCAAGGCACAGCAAGATAACTTTCCACTGGGGAAAATGCAATGTGTCACTCTGAGGATCG	525
Qy	141	GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTrpCysPheAaspAla	160
Db	526	GAACATAGCTTAGAAGNCAGAGTTCCTCACTGGAGATGCAAGTCTACTGCTTTGATCGG	585
Qy	161	AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer	180
Db	586	GACAGATTTTCCAGTTTGGAGACGAGTTAAAGGAAAAGAAAGATTAAGGCGTTTATCC	645
Qy	181	IleLeuPheGluValGlyThrGluGluAenLeuAaspPheLysAlaIleIleAaspGlyVal	200
Db	646	ATTTTATTTGAGGTTGGAGTTTGAAGAAAATTTTGGATTACAAAGGCATTTATTTGATGGA	705
Qy	201	GluSerValSerArgPheGlyLysGlnAlaAlaLeuAaspPropheIleLeuLeuAenLeu	220
Db	706	GAGAGTGTAGTCTGTTTGGAAAGCAGCGTCTTTAGATCCATTCGCTTCGTCAGAACCTC	765
Qy	221	LeuProAenSerThrAaspLysTrpTrpIleTrpAenGlySerLeuThrSerProProCys	240
Db	766	CTGCAAAACTCCACTGACAAGTATTACATTTACAATGGATCATTTGACATCCCTCCCTGC	825
Qy	241	ThrAaspThrValAaspTrpIleValPheLysAaspThrValSerIleSerGluSerGlnLeu	260
Db	826	ACACACACCCGGGAATGGATTGTTTTTAAGGATACAGTTAGCATCTCTGAAAAGCCAGCTG	885
Qy	261	AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTrpValMetLeuMetAaspTrp	280
Db	886	GCTGTATTTGTGAAGTCTCACAAATGCACAGCTCGGGTATGTCATGTTGATGGATTAC	945
Qy	281	LeuGlnAenAenPheArgGluGlnGlnTrpLysPheSerArgGlnValPheSerSerTrp	300
Db	946	TTACAAAACAATTTCCGAGAACACACAGTACAAAGTTTTCCAGCGAGGTGTTTTCTCATAT	1005
Qy	301	ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAenValGlnAla	320
Db	1006	ACTGGAAAGGAAGAGATCCACGAAAGTAGTGTGTAGTTTCAGAACCCAGAAAATGTGCAAG	1065
Qy	321	AaspProGluAenTrpThrSerLeuLeuValThrTrpGluArgProAaspValValTrpAasp	340
Db	1066	GACCTGTGAATTAACACAGCCTTCTGTGTCAATGGGAAAGACCTTCGGGTGTTTATGAC	1125
Qy	341	ThrMetIleGluLysPheAlaValLeuTrpGlnGlnLeuAaspGlyGluAaspGlnThrLys	360
Db	1126	GCCATGATTGAGAAGTTTGCAGTCTCTGTACAGCCACTGGCGGGAAAATGACCAAGCCAAAG	1185
Qy	361	HisGluPheLeuThrAaspGlyTrpGlnAaspLeuValThrIle	374
Db	1186	CATGAGTCTTAAACAGATGGCTATCAGGACTTGGGTGCCATT	1227
RESULT 14			
NMU428208			
LOCUS			
DEFINITION Mus musculus mRNA for phosphacan short isoform (RPTP-beta gene).			
ACCESSION AJ428208			
VERSION AJ428208.1 GI:29465987			
MMU428208 3784 bp mRNA linear ROD 24-JUN-2003			

RESULT 14
MMU428208
LOCUS
DEFINITION
ACCESSION
VERSION

Db 61 TGGGCTATGGTACTACAGACAGAGAACTTGTGAAGAGATTGGCTGGTCTTAC 120
QY 41 ThrGluValLeuAenGlnLysAenTTPGlyLysLysTyrProThrCysAenSerProLys 60
Db 121 ACAGAGCACTAAATCAAAAATTTGGGAAGAATATCCATGTGTGAATAGCCCAAG 180
QY 61 GlnSerProileAenLysLeuAenGlnLysLeuThrGlnValAenValAenLysLysLeu 80
Db 181 CAGTCTCTTATTAATTCACGAGATCTTACACAGTCAATGTGAATCTTAAGAAGCTG 240
QY 81 LysPheGlnGlyTTPaspLysThrSerLeuGluAenThrPheIleHisAenThrGlyLys 100
Db 241 AAATTTCCAGGGTTGGAAAAGCGTCTTGGAAAACACCGTTTCAATTCACAGCACTGGAAA 300
QY 101 ThrValGluLeuAenLysLeuThrAenSerPheValSerGlyValSerGluMetVal 120
Db 301 ACAGTGGAAATTAATCTCAATTAATGACTACTATCTACGTTGGAGGACTTTTCAGAAAAGGTC 360
QY 121 PheLysAlaSerLysIleThrPheHisTTPGlyLysCysAenMetSerSerAspGlySer 140
Db 361 TTCAAGGCAAGCAAGATACTTTCCACCGGGGAAAATGCAATGTGTCTATCTGAAGATCG 420
QY 141 GluHisSerLeuGlnGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 421 GAACATAGCTTAGAAGGACAGAAAGTCCCACTGGAGATGCAAGTCTACTGCTTTGATGCG 480
QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 481 GACAGATTTTCAGTGTTCAGATGAGGAGCGTTAAAGGAAAAGAAAGATTAAGGCGCTTTATCC 540
QY 181 IleLeuPheGluValGlyThrGluGluAenLeuAenPheLysAlaIleLeuAspGlyVal 200
Db 541 ATTTTATTTGAGTTGGAGTTCAGAGAAATTTGGATTACAAAGCCATTTATGATGGAAT 600
QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAenPheProPheIleLeuAenLeu 220
Db 601 CAGAGTGTGTAGTGTGTGGAAAAGCAGGCTGCTTAGATCCATTCCTGCTTGGCAACCTC 660
QY 221 LeuProAenSerThrAspLysTyrTyrIleTyrAenGlySerLeuThrSerProCys 240
Db 661 CTGCAAAACTCCACTGACAGATTTACATTTTACAAATGATCATTCACATCCCTCCCTGC 720
QY 241 ThrAspThrValAspTTPileValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 721 ACAGACACCGTGAATGGATGTTTAAAGGATACAGTTAGCATCTCTGAAGCCAGCTG 780
QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 781 GCTGTATTTTGTGAAGTTCTCAATGCAACAGTCTGGGTATGTCTATGTTGATGATTAC 840
QY 281 LeuGlnAenAenPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 841 TTACAAAACAAATTTCCGAGAACACAGTACAGTATTTCCAGGCAGGTGTTTTCTTCATAT 900
QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAenValGlnAla 320
Db 901 ACTGAAAGAAAGATCCACGAGTAGTGTGTGTTTACAAACCAAGAAATGTGCAACT 960
QY 321 AspProGluAenTyrThrSerLeuLeuValThrTTPGluAenArgProArgValValTyrAsp 340
Db 961 GACCTGTAGAAATTAACACAGCCCTCTGTGTACATCGGAAAGACCTCGGGTGTGTTATGAC 1020
QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAenPheGlyLysGlnThrLys 360
Db 1021 GCCATGATGGAAGATTTTCAGTCTCTGACCAAGCCACTGGCGGGAATGACCAAGCCCAAG 1080
QY 361 HisGluPheLeuThrAspGlyTyrGlnAenLeuValThrIle 374
Db 1081 CATGAGTTCTTAACAGATGGCTATCAGGACTTGGGTGCCATT 1122

RESULT 15
BC070850

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
JOURNAL
REMARK
COMMENT

BC070850 5281 bp mRNA linear VRT 19-JUL-2004
Xenopus laevis hypothetical protein MGC84589, mRNA (cdna clone
MGC:84589 IMAGE:6640110), complete cds.
BC070850
BC070850.1 GI:47682857
MGC.
Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 5281)
Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
and Richardson,P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative
Dev. Dyn. 225 (4), 384-391 (2002)
12454917
2 (bases 1 to 5281)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,P.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Carniez,T.E., Brownstein,M.J., Udén,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaney,S.J., Bosak,S.A., McSwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Binkley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
Schnur,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
3 (bases 1 to 5281)
Klein,S. and Strausberg,R.
Direct Submission
Submitted (10-MAY-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: cagpb-remail.nih.gov
Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu,
Parvaneh Saeedi, JR Santos, Angélique Schnerch, Ursula Skaleka,
Duane Smalish, Jeff Stott, Miranda Tsai, George Yang, Jacquele
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 158 Row: d Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers

```
source      1. .5281
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:84589 IMAGE:6640110"
/tissue type="Oocytes"
/clone_lib="NICHD_XGC_001"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6.1"
1. .5281
gene
/gene="MGC84589"
/db_xref="LocusID:432027"
187_ .4560
/gene="MGC84589"
/codon_start=1
/product="MGC84589 protein"
/protein_id="AAH70850.1"
/db_xref="GI:4768288"
/translators="MLNPRCLVCLQFLPSVMQVGLPFSYVRPQRKLTETIDMSVTETL
NQRWAKYIPACNGAKSPINIDEDLTQNLNKGHFEQWDEKSSSYTLIRNKGTV
EININYYLSGALDSQYKASRITFHGNCNATADGSEHSLEGQKPLEMQIFCFNH
NEFNIDAEISNGKIKALSILFKVQGEDNRYNAI IDGVNVRHFGTKVLAEPFELS
NLDPDYDKYRHFSTGPPSTVEWIVFKDTVTISESOLEMFCELTWQSGYVM
VMDLYNNRHRQEFSGQVFSSTGTETIEIPKVCSEPEDEVRADSONNTSLITWR
PRVYDAVIERLYVYQHLEGEDQTKHEHLTDGYQDILGALLIYNLPNTSYALQVLAIC
TNGLYGLKSDQYIVDMTLDPOGAPFDITDFTETEDIQSESEVEEKENESRKSLSV
SNTNQKQAGQQAATSSYSEGNAYLTGKTADDKLRHVDVLAAYLPTLQTTKTMH
LTDVYSGKASGALDQSPQSEATNLSKOPKSRVEIWMNLNIDSTKSIABEDQTSI
DSELOQTTNKKRLLIIPRSPKPVNSHTSVSIPTAIIYSSEITSGQNI VSNIPDSI
DDAFISIEDISKVTAPTREPLDVTIISITSKSNVTGEGHPTNSINLLKDELLSTSF
SSGATVQATASSLEDVYNKATLPQEMPLVPNPFSTPEVLLESVSGAGVGHSESSV
SEVUSQTALSTTPVNSAENSHESRVGLAESEEEKTVIPLVVSSLPFICLVV
LVGLIYRNECFKFAHPEDKTSRISGPIIPFPVADVDDVAGVLPKPKHVEDLH
SRNGFTDEFKLEQYQIQCNTADLGTISENSNHPNKNQRYVNTVAYDHSRVDLH
QLAEKGRDQDYINANFVGNKQYIAAQKPLKTAEDFWRMIWEHNEVEI VMI TN
LVEGRKCDQYIPSDNGEYGA LVTKQKTVLA VYTRFTFVRNKKIKKQSQKGRQ
NERAVQYHYTQFDMGVPEYTLPLVTFVRKASATKNGPVPVHCSAGVGRGTYYIVL
DSMLQQLIHGETVNIIFGLHRSQRNLYVOTEQYIPIHDALVEAILLSRETAVPEYTH
IHGYVNTLLTTPSGKRLSTQFKLSEPNLLQCDYSTALKPQNKDKRTSSVIPER
SRVGLSLAGSGDTYINASYVMGYHESNQFIITQHPLLHTIKDFWRMMWDHNSQFIYM
LPGMNMWAEDEBFWPKNQDPIVCDTFTVLVIGRHDVCLNSNEEMIYQDFILEATODD
YVLRVRFQCPKPNPDS PMSKTFELINLKEBAANKGTIIIVHDEYGGVTAGTFCAL
YRLMHQLEQENADVQYQVAXMINLMRPGIFDIDIEQYQFLYKAMLSLVSTRQEEDSMVP
MDTNGAFEDPDGSTAESLESIV"
ORIGIN
Alignment Scores:
Pred. No.:      2.83e-124      Length:      5281
Score:          1404.00      Matches:      255
Percent Similarity: 83.01%      Conservative: 48
Best Local Similarity: 69.86%      Mismatches: 62
Query Match:      70.91%      Indels:      0
DB:              5          Gaps:      0
US-09-983-000a-2 (1-374) x BC070850 (1-5281)
Qy      7 PheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAspTrpAlaAsnGlyTyrTyr 26
Db      205 TTTCTGGTGTGCTTCAGCTCTTTTCTGTCGCAAGTGGGTTTGGCGGTTTTCATATTAC 264
Qy      27 ArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyrThrGlyAlaLeuAsnGln 46
Db      265 AGACCAACAGCGTAATAATTAACGTGAAGAATAAGATTGGTCTTACACAGAAACCTTAACCAA 324
Qy      47 LysAsnTrpGlyLysLysTyrProThrCysAsnSerProLysGlnSerProIleAsnIle 66
Db      325 GAAAACTGGGCAAAAAAATAATCCAGCATGCAATGGTGCACCAACCAATCTCCTATCAACATT 384
Qy      67 AspGluAspLeuThrGlnValAsnValAsnValAsnLeuLysLysLeuLysPheGlnGlyTrpAsp 86
Db      385 GATGAGGATTTAAACACAAATTAACACAAATTTAAAGAAATTTACATTTTGGGGTTGGGAT 444
```

```
Qy      87 LysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLysThrValGluIleAsnLeu 106
Db      445 AAGAGTCCTCGGGAATATACACTAATTCGAAAAATCTGAAAAACAGTGGAAATCAATTTG 504
Qy      107 ThrAsnAspTyrArgValSerGlyValValSerGluMetValPheLysAlaSerLysIle 126
Db      505 ACAAAATGATTATTATTGAGTGGTGGTGGCTTAGACTCACAAATACAAAGCAAGCAGAATA 564
Qy      127 ThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySerGluHisSerLeuGluGly 146
Db      565 ACGTTTCATTTGGGGAAACTGCAATGCTACAGCTGATGCTTCCGAAACACAGCTTAGAAGGC 624
Qy      147 GlnLysPheProLeuGluMetGlnIleTyrCysPheAspAlaAspArgPheSerSerPhe 166
Db      625 CAGAAGTTCCTTTAGAGATGCAAAATCTTCTGCTTTTAATCATACGAGTTTAAAAATATT 684
Qy      167 GluGluAlaValLysLysGlyLysLeuArgAlaLeuSerIleLeuPheGluValGly 186
Db      685 GACGAACCAATAAGTGGAAATGGCAAAATAAAAGCTTTGTCCATTTTATTAAAGGTTGGC 744
Qy      187 ThrGluGlnAsnLeuAspPheLysAlaIleIleAspGlyValGluSerValSerArgPhe 206
Db      745 CAGGAAGATAACGTCGCTTATAATGCAATCATCGATGAGTAGATAATGATCATCGTTTT 804
Qy      207 GlyLysGlnAlaAlaLeuAspProPheIleLeuLeuAsnLeuLeuProAsnSerThrAsp 226
Db      805 GGTACAAAAGTTGCAITGGAGGCCATTTTCTGTCAAAATCTTCTCTGACTACACCGAC 864
Qy      227 LysTyrTyrIleTyrAsnGlySerLeuThrSerProCysThrAspThrValAspTrp 246
Db      865 AATATTATATACATATAATGCTCACTAACCAACCCCTCCGTGCTCAGAAACACAGTGAATGG 924
Qy      247 IleValPheLysAspThrValSerIleSerGluSerGlnLeuAlaValPheCysGluVal 266
Db      925 ATTGTGTTTAAAGACACACAGTTACAAATTTCAAGAACCCAGTTGGAAATGTTTGTGAGATC 984
Qy      267 LeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyrLeuGlnAsnAsnPheArg 286
Db      985 CTTACAATGAGCAATCTGGGTATGTCATGGTGAATGATTTTGGAAAAACAACATATAGA 1044
Qy      287 GluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyrThrGlyLysGluGluIle 306
Db      1045 CACCAGCAGTTTGAGTTCTCAGGACCAAGTATTTTTCATCTTATACTGTCACAGAGGAGATC 1104
Qy      307 HisGluAlaValCysSerSerGluProGluAsnValGlnAlaAspProGluAsnTyrThr 326
Db      1105 CCGGAACCTGTTTGCAGCTCTGAACCCAGAGAGCGTCCGAGCTGATTCCACAGAAATAACACA 1164
Qy      327 SerLeuLeuValThrTrpGluArgProArgValValTyrAspThrMetIleGluLysPhe 346
Db      1165 AGCTTCTCTACATACATGGAGAGACCACCGGGTGTGTTTATGATGTCAGTGTAGAGAGATAC 1224
Qy      347 AlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLysHisGluPheLeuThrAsp 366
Db      1225 CTGGTCTTCTCAACATTTTAGAAGGAGAGAAGTCAAACTAAGCATGAACATCTCACTGAT 1284
Qy      367 GlyTyrGlnAspLeu 371
Db      1285 GGTATCAAGACCTG 1299
```

Search completed: August 30, 2005, 11:10:37
Job time : 4363 secs

Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Tanaka, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (KISA) system--384-format
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE PUBMED AUTHORS

11076861
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection

JOURNAL REFERENCE AUTHORS

Nature 409, 685-690 (2001)
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.

JOURNAL REFERENCE AUTHORS

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

6 (bases 1 to 1614)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

JOURNAL

Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp,
URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216]

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>

FEATURES

source

1..1614
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:C130068022"
/db_xref="taxon:10090"
/clone="C130068022"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="16 days embryo"
327..>1613
/note="unnamed protein product; protein tyrosine
phosphatase, receptor type, Z (MGD)GI:97816
GB|NW_011219, evidence: BLASTN, 99%, match=1393)
putative"
/codon_start=1
/protein_id="BAC38299.1"
/db_xref="GI:26349319"

CDS

translation="WRILQSLFACVOLLCLRLDWAYGYRQORQLVEEIGSVTGAL
NOKWGEKYPICNSPKQSPINIDELTQVNVNKLKQCKEASLENTPIHNTGTV
EINTNDYLSGGUSEKVFVKASKITFMKGNVNSSEGSLEQKQKPLEMQVYCPDA
DRFSFEAVKGRGLRALSILFVGVVEENLDYKALIDGTSVSRFGKQALDPVLIQ
NLLFNSTDKYIYNGSLTSPCTVEMIVFKDVTLSIESQLAFCEVLTMQOQSYVM
LMDYIQQNFRQYKFSRQVFSYTGKEIEHVVCSSSEPENQADPNVTSLLVTVWR
PRVYVDAMIEKFAVLQPLAGNDQAKHFLTDGVDLQILNLLPNNSVYLQIVAVC

SNGLYGYKSPQLIVDMPTEDAELELDPPELIGTEEEK"

ORIGIN

Alignment Scores:

Pred. No.: 8,85e-208 Length: 1614
Score: 1844.00 Matches: 344
Percent Similarity: 95.45% Conservative: 13
Best Local Similarity: 91.98% Mismatches: 17
Query Match: 93.13% Indels: 0
DB: 3 Gaps: 0

US-09-983-000A-2 (1-374) x AK081698 (1-1614)

Qy 1 MetArglleLeuLysArgPheLeuAlaCyslleGlnLeuLeuCysValCysArgLeuAsp 20

Db 327 ATGCGAATCTCGCAGAGCTTCTCGCGTGGTTCAGCTCTGCTGCTGCGCTCGAC 386

Qy 21 TTPAlaAAsnGLYTYrTYrArgGlnGlnArgLysLeuValGluGluIleGlyTTPSerTyr 40

Db 387 TGGCTTATGGATACTACAGACAACAGAGAAACTTGTTCAGAGAGATTGGCTGGCTCTAC 446

Qy 41 ThrGlyAlaLeuAAsnGlnLysAAsnTTPGlyLysLysTyrProThrCysAAsnSerProLys 60

Db 447 ACAGGAGCACTAAATCAAAAAAATTTGGGAAAGAAATATCCAATATGTAATAGCCCAAG 506

Qy 61 GlnSerProIleAAsnIleAspGluAAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80

Db 507 CAGTCTCTTATTAATATTTGATGAAGATCTTACCAAGTCAATGGAATCTTTAAGAAATCG 566

Qy 81 LysPheGlnGlyTTPAspLysThrSerLeuGluAAsnThrPheIleHisAAsnThrGlyLys 100

Db 567 AATTTTCAGGGTTGGGAAAAAGCGTCTTGGAAAAACAGCTTCATTCACACACTGGGAAA 626

Qy 101 ThrValGluLeuAAsnLeuThrAAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120

Db 627 ACAGTGAATTAATCTCACTAATGACTATCTCAGTGGAGGACTTTCAGAAAAAGGTC 686

Qy 121 PheLysAlaSerLysIleThrPheHisTTPGlyLysCysAAsnMetSerSerAspGlySer 140

Db 687 TTCAGGCAAGCAAGATAACTTTCCACTGGGAAAAATGCAATGTGTCTATCTGAAGATCG 746

Qy 141 GluHisSerLeuGluGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160

Db 747 GAACATAGCTTAGNAGACAGAGAGTTCCTCACTGGAGATGCAAGTCTACTGCTTTGATCG 806

Qy 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180

Db 807 GACAGATTTTCCAGTTTTCAGGAAGCAGTTAAAGGAAAAAGAGATTTAAGGGCTTTATCC 866

Qy 181 IleLeuPheGluValGlyThrGluGluAAsnLeuAspPheLysAlaIleLeuAspGlyVal 200

Db 867 ATTTTATTTGAGGTTGGAGTTGAAGAAAAATTTGGATTACAAGCCATTTATTGATGGAAC 926

Qy 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAAsnLeu 220

Db 927 GAGAGTGTAGTCGTTTGGAAAGCAGCGTCTTTAGATCCATTCGTTGTCAGAACCTC 986

Qy 221 LeuProAAsnSerThrAspLysTyrTyrIleTyrAAsnGlySerLeuThrSerProProCys 240

Db 987 CTGCCAAATCCCACTGACAAGTATTACATTTACAATGGATCATTTGACATCCCTCCCTCGC 1046

Qy 241 ThrAspThrValAspTTPIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260

Db 1047 ACAGACACCGTGGAAATGGATTTGTTTTTAAGGATACAGTTAGCATCTCTGAAGGCCAGCTG 1106

Qy 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280

Db 1107 GCTGATTTTGTGAAGTCTTCACAACTGCAAGCTCTGGGTATGTCATGTTGATGATTAC 1166

Qy 281 LeuGlnAAsnAAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300

Db 1167 TTCAAAAAAATTTCCGAGAACCAACAGTACAAGTTTTCCAGGCCAGGTGTTTTCTCTCATAT 1226

QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerSerGluProGluAsnValGlnAla 320
 Db 1227 ACTGGAAGAGAGAGATCCAGAGTAGTGTAGTTCAGAACACAGAAATGTGCAAGCT 1286
 QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTTPGluAatGProArgValValTyrAsp 340
 Db 1287 GACCTTGAGATTACACAGACCTTCGTGTCATCGGAAGACCTCGGGTGTGTTATGAC 1346
 QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
 Db 1347 GCATGATTGAGAAGTTTCAGTTCGTACACAGCCTGGCGGGAATGACCAAGCCCAAG 1406
 QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
 Db 1407 CATGAGTTCTTAACAGATGGCTATCAGGACTTGGGTGCCATT 1448

RESULT 2
 CD350667
 LOCUS
 DEFINITION UI-M-G10-cgg-f-21-0-UI.r1 NIH BMAP_G10 Mus musculus cDNA clone
 IMAGE:6852982 5', mRNA sequence.
 ACCESSION CD350667
 VERSION CD350667.1 GI:31142182
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 757)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. 757
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6852982"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5, 14.5, 16.5, 17.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP_G10"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonafo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 is AGCGAGACAG. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Alignment Scores:

Pred. No.:

7.25e-135 Length:

757

Score: 1227.00 Matches: 232
 Percent Similarity: 97.19% Conservative: 10
 Best Local Similarity: 93.17% Mismatches: 7
 Query Match: 61.97% Indels: 0
 DB: 6 Gaps: 0

US-09-983-000A-2 (1-374) x CD350667 (1-757)

QY 58 SerProLysGlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeu 77
 Db 3 AGCCCAAGAGAGTCTCTCTATTATATATATATATATATATATATATATATATATATAT 62
 QY 78 LysLysLeuLysPheGlnGlyTTPAspLysThrSerLeuGluAsnThrPheIleHisAsn 97
 Db 63 AAGAAACTGAAATTTTCAGGGTTGGGAAAGCGTCTTGGGAAACACGTTCAATTCACAAC 122
 QY 98 ThrGlyLysThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSer 117
 Db 123 ACTGGGAAACAGTGGGAAATAAATCTCATAATCACTACTATCTCAGTGGAGGACTTTCA 182
 QY 118 GluMetValPheLysAlaSerLysIleThrPheHisTTPGlyLysCysAsnMetSerSer 137
 Db 183 GAAAGAGTCTTCAAGGCAAGCAGATTAATTTCCACTGGGGAATAATGCATGTGTCTCT 242
 QY 138 AspGlySerGluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCys 157
 Db 243 GAAGGATCGGAACATAGCTTAGAAGGACAGAAGTTCCCACTGGAGATGCAAGTCTACTGC 302
 QY 158 PheAspAlaAspArgPheSerSerPheGluGluAlaValLysGlyLysLeuArg 177
 Db 303 TTTGATCGGACAGATTTTCCAGTTTTCAGGAGACAGTAAAGGAAAGGAAGATTAAAG 362
 QY 178 AlalaSerIleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIle 197
 Db 363 GCATTATCCATTTTATTTGAGTTTGGAGTTGAAGAAATTTGGATTACAAAGCCATTATT 422
 QY 198 AspGlyValGluSerValSerArgPheGlyGlnAlaAlaLeuAspPropheIleLeu 217
 Db 423 GATGGAACCTGAGAGTGTAGTTCGTTTGGAAAGCAGGCTGCTTTAGATCCATTCTCTGA 482
 QY 218 LeuAsnLeuLeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSer 237
 Db 483 CAGAACCTCTCTGCCAAATCCCACTGACAGATTTATACATTTTACATGATGATTCATCC 542
 QY 238 ProProCysThrAspThrValAspTTPileValPheLysAspThrValSerIleSerGlu 257
 Db 543 CTTCCCTGCAACAGACACCGTGAATGATTTTAAAGGATACAGTTAGCATCTCTGAA 602
 QY 258 SerGlnLeuAlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeu 277
 Db 603 AGCCAGCTGGCTGTATTTTGTGAAGTTCTCACAATGCAACAGCTGGGTATGTCATGTTG 662
 QY 278 MetAspTyrLeuGlnAsnAsnAspArgGluGlnGlnTyrLysPheSerArgGlnValPhe 297
 Db 663 ATGGAATTACTTACANACAAATTTCCGAGAACACACAGTACAAAGTTTTCAGGACAGGT 722
 QY 298 SerSerTyrThrGlyLysGluGluIle 306
 Db 723 TCCTCATATATCTGGAAGGAAGAGATC 749

RESULT 3

BG911384

LOCUS 755 bp mRNA linear EST 05-JUN-2001

DEFINITION 620807687F1 NCI CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4939965

5', mRNA sequence.

ACCESSION BG911384

VERSION BG911384.1 GI:14291860

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 755)

AUTHORS TITLE JOURNAL COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC Clone Distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM10877 row: m column: 22
High quality sequence stop: 747.
Location/Qualifiers
1. 755
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4939965"
/tissue_type="anaplastic oligodendroglioma with lp/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Brn67"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

FEATURES source

ORIGIN

Alignment Scores:
Pred. No.: 2,186-123 Length: 755
Score: 1130.50 Matches: 242
Percent Similarity: 96.05% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 10
Query Match: 57.10% Indels: 6
DB: 4 Gaps: 0

US-09-983-000a-2 (1-374) x BG911384 (1-755)

QY 52 LysfyrProThrCysAsnSerProlysglnSerProileAsnileAspGluAspLeuThr 71
DB 2 AAATATCAACATGTAATAGCCCAACAACTCTCTCATATAT-GATGAAGATCTTACA 60
QY 72 GlnValAsnValAsnLeuLysLeuLysPheGlnGlyTrpAspLysThrSerLeuGlu 91
DB 61 CAAGTAAATGTGAATCTTTAAGAAACTTAAATTTTCAGGCTTGGGATAAAACATCATG-GAA 119
QY 92 AsnThrPheIleHisAsnThrGlyLysThrValGluIleAsnLeuThrAsnAspTyrArg 111
DB 120 AACACATTCATTAACATCGGGAACAGTGGMAATTAATCTCACTAATGACTACCGT 179
QY 112 ValSerGlyValSerGluMetValPheLysAlaSerLysIleThrPheHisTrpGly 131
DB 180 GTCACGGAGGAGTTTCAGAAATGGTGTTTAAGCAAGCAAGATAACTTTTCACTGGGGA 239
QY 132 LysCysAsnMetSerSerAspGlySerGluHisSerLeuGlyGlnLysPheProLeu 151
DB 240 AAATGCAATATGTCACTCTGATGGATCAGAGCATATGTTTAGAAGGACAAATAATTTCCACTT 299
QY 152 GluMetGlnIleTyrCysPheAspAlaAspArgPheSerSerPheGluGluAlaValLys 171
DB 300 GAGATGCAAACTACTGCTTTGATGGGACCGCATTTTTCAG-TTTGAGGAGCGAGTCMAA 358
QY 172 GlyLysGlyLysLeuArgAlaLeuSerIleLeuPheGluValGlyThrGluAsnLeu 191
DB 359 GGAAAAGGGAAGTTTAAGAGCTTTATCCAT-TTGTGTTGAGGTTGGACAGAAATAATTTG 417
QY 192 AspPheLysAlaIleIleAspGlyValGluSerValSerArgPheGlyLysGluAlaAla 211
DB 418 GATTTTCAAAAGCGATTAATGATGGATCGAAAGTGTAGTCGTTCTGGGAACAGCGCTGCT 477
QY 212 LeuAspProPheIleLeuAsnLeuLeuProAsnSerThrAspLysTyrTyrIleTyr 231

DB 478 TTAGATCCATTATACATCTGTTGAACCTTCTGCCAAACTCACTGCAAGTATTACATTAC 537
QY 232 AsnGlySerLeuThrSerProCysThrAspThrValAspTrpIleValPheLysAsp 251
DB 538 AATGGCTCAATTGACATCTCTCCCTGCACAGACACAGTTCGACTGATGTTGTTTAAAGAT 597
QY 252 ThrValSerIleSerGluSerGlnLeuAlaValPheCysGluValLeuThrMetGlnGln 271
DB 598 ACAGTTAGCATCTCTGAAGCCAGTGGCTGTTTA-TGTGAAGTTCCTTACATGCAACAA 656
QY 272 SerGlyTyrValMetLeuMetAspTyrLeuGlnAsnAsnPheArgGluGlnGlnTyrLys 291
DB 657 TCTGGTTATGTCATGCTGATGGACTACTTACAAAAGAAATTTTCGAGAGCACAGTACAAG 716
QY 292 PheSerArgGlnValPheSerSerTyrThrGlyLysGlu 304
DB 717 TTCTCTAGACAGGTGTG-TTCTCATACACTGGACAGGAA 754

RESULT 4 CN383698 LOCUS

DEFINITION 701 bp mRNA linear EST 01-OCT-2004
4117208 BARC 3GAL chicken mixed tissue Gallus gallus cDNA clone
3GAL 9B14 5', mRNA sequence.

ACCESSION CN383698
VERSION CN383698.1 GI:53550459
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

REFERENCE Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archoosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 701)

AUTHORS Evoek-Clover,C.M., Ashwell,C.M., McMurtry,J.P., Lillehoj,H.S.,
Matukumalli,L.K. and Van Tassel,C.P.

TITLE Characterization of expressed sequence tags generated from multiple
chicken tissues
Unpublished (2004)

COMMENT Contact: Christina M. Clover
Growth Biology Laboratory
Animal and Natural Resources Institute
Bldg. 200 Rm215 BARC-East, Beltsville, MD 20705, USA
Tel: 3015048224
Fax: 3015048623

Email: chrisc@nri.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim alt -trim fastaVector identified by
cross match using options -minmatch 12 -minscore 12

Plate: 9 row: B column: 14

Seq primer: CCCAGTCACGACGTTGTAAAACG

High quality sequence stop: 701.

FEATURES source

Location/Qualifiers
1. 701
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Leghorn and broiler"
/db_xref="taxon:9031"
/clone="3GAL_9B14"
/lab_host="DH5alpha"
/note="Vector: pBluescript, SK+, Stratagene; Site 1: NotI;
Site 2: EcoRI; Normalized library of pooled RNA isolated
from whole brain, ultimobranchial gland, parathyroid
gland, cecal tonsil and primordial germ cells Multiple"

ORIGIN

Alignment Scores:
Pred. No.: 7e-116 Length: 701
Score: 1067.00 Matches: 196
Percent Similarity: 94.37% Conservative: 22
Best Local Similarity: 84.85% Mismatches: 13
Query Match: 53.89% Indels: 0
DB: 7 Gaps: 0

US-09-983-000A-2 (1-374) x CN383698 (1-701)

```

QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 4 GAACATAGCTTGAAGGACAAATTTCTCTGAGATGCAATCTACTGCTATGGA 63
QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 64 GATCTGTTTACAGACTTCAAGAGGCAATTAAGGAATGGAAGTTAAGAGCTTATCA 123
QY 181 IleLeuPheGluValGlyThrGluGluLeuLeuAspPheLysAlaIleLeuAspGlyVal 200
Db 124 GTTGTGTTGAGATTGAGTAGAAGATATCCGATTAATTTCCATCAATTAAGGAGTA 183
QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuLeuLeu 220
Db 184 GATAGTGTGTTGCTGTTGGAANAACAGCTGCTTAGAACCAATTTGTTGCTGAACCTT 243
QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProCys 240
Db 244 TTACCGAATGCAACGACCAATATTACACTTACATGGTCTTTATCAGCTCCTCTGC 303
QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 304 TCAGAAACAGTTGAATGGAATGCTGTTCAAGATACATATTAGTATTCTGAGCAACAGTTA 363
QY 261 AlaValPheGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 364 GCAGTATTCTGGAAGTCTCTTACAATGCAACAGTCTGGCTATGCTATGCTATGCTATG 423
QY 281 LeuGlnAsnAsnPheArgGluGlnGlnIleTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 424 CTGCAAAACACTTTCGAGAGCAACGATATAGTTCTCTGGCAAGTGTCTTCTCTTAC 483
QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 484 ACTGCAGACGAGAGAAATTCATGACGAGTTTGCACTGAGACCTGAAATGTCCAATCT 543
QY 321 AspProGluAsnTyrSerLeuLeuValThrTyrGluArgProArgValValTyrAsp 340
Db 544 GATCCAAAGAAATTAACAGACCTCTCTGTTACGTGGGAAAGACCTCGGGTGTGTATGAT 603
QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 604 ACCATGATTGAGAAGTTGCTGCTTTTATCAACAGTTGGATGGAGAGACCAAGCTAAG 663
QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeu 371
Db 664 CATGAGTTTCTGACAGATGGGTATCAGGACTTG 696

```

```

RESULT 5
CN418533 572 bp mRNA linear EST 16-MAY-2004
LOCUS 17000424334377 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN418533
VERSION CN418533.1 GI:47406127
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 572)
Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J. and Stanton,L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA

```

Tel: 650 473 8658
 Fax: 650 473 7760
 Email: xbrandenberger@geron.com
 Insert Length: 572 Std Error: 0.00.

FEATURES

Location/Qualifiers
 1..572
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, cell lines H1, H7, and H9"
 /clone_lib="GRN_ES"
 /notes="Oligo dt primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

ORIGIN

Alignment Scores:
 Pred. No.: 1,29e-106 Length: 572
 Score: 988.00 Matches: 187
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.47% Mismatches: 0
 Query Match: 49.90% Indels: 0
 DB: 7 Gaps: 0

US-09-983-000A-2 (1-374) x CN418533 (1-572)

```

QY 29 GlnArgLysLeuValGluGluIleGlyTyrSerTyrThrGlyAlaLeuAsnGlnLysAsn 48
Db 7 AGGAGAAACCTTGTGAAGAGATTGGCTGCTCTATACAGGAGCACTGAATCAAAAAAT 66
QY 49 TrpGlyLysLysTyrProThrCysAsnSerProLysGlnSerProIleAsnIleAspGlu 68
Db 67 TGGGGAAAGAAATATCCACATGTAAAGCCCAAAACATCTCTATCAATATTGATGAA 126
QY 69 AspLeuThrGlnValAsnValAsnLeuLysLysLysPheGlnGlyTyrTrpAspLysThr 88
Db 127 GATCTTACACAAGTAATATGTGAATCTTAAAGAAACTTAAATTTCAAGGTTGGGATAAACA 186
QY 89 SerLeuGluAsnThrPheIleHisAsnThrGlyLysThrValGluIleAsnLeuThrAsn 108
Db 187 TCATTGGAAACACATTCATTAACACTGGGAAACAGTGGAAATTAATCTCACTAAT 246
QY 109 AspTyrArgValSerGlyValSerGluMetValPheLysAlaSerLysIleThrPhe 128
Db 247 GACTACCGTGTACGCGAGGAGTTTCAGAAATGGTGTGTTAAGCAAGCAAGTAATCTTT 306
QY 129 HistTrpGlyLysCysAsnMetSerSerAspGlySerGluHisSerLeuGluGlyGlnLys 148
Db 307 CACTGGGGAAAAATGCAATATGTCTATGTGATCAGAGCATAGTTTAGAAGGACAAAAA 366
QY 149 PheProLeuGluMetGlnIleTyrCysPheAspAlaAspArgPheSerSerPheGluGlu 168
Db 367 TTTCCACTTGGAGATGCAAAATCTACTGCTGTGATGTCAGACCGATTTTCAAGTTTGGAGAA 426
QY 169 AlaValLysGlyLysGlyLysLeuArgAlaLeuSerIleLeuPheGluValGlyThrGlu 188
Db 427 GCAGTCAAGAGAAAGGAAAGTTAAGAGCTTTATCCATTTTGTGAGGTTGGGACAGAA 486
QY 189 GluAsnLeuAspPheLysAlaIleIleAspGlyValGluSerValSerArgPheGlyLys 208
Db 487 GAAAAATTTGGATTTCAAGGCGATTTATGATGGAGTCGAAAGTGTAGTCTGTTTGGGAGAG 546
QY 209 GlnAlaLeuAspProPheIle 216
Db 547 CAGGCTGCTTTAGATCCATTCAATA 570

```

RESULT 6

```

CA879683 585 bp mRNA linear EST 20-DEC-2002
LOCUS K0976A09-5N NIA Mouse Neural Stem Cell (Undifferentiated) cDNA
DEFINITION Library (Long) Mus musculus cDNA clone NIA:K0976A09 IMAGE:30091496
5', mRNA sequence.

```


ACCESSION CA879683
 VERSION CA879683.1 GI:27331232
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM
 REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 TITLE Piao, Y., Dudekula, D.B., Qian, Y., Martin, P.R., Alba, K., Vescovi, A.L. Systematic Analyses of NIA Mouse Neural Stem Cell (Undifferentiated) cDNA Library (Long)
 JOURNAL Unpublished (2002)
 COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@gsun.grc.nia.nih.gov
 Plate: K0976 row: A column: 09
 Seq primer: -21M13 Reverse
 High quality sequence stop: 585
 POLYA=No.

FEATURES
 source
 1..585
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CD1"
 /db_xref="niaEST:K0976A09-5N"
 /db_xref="taxon:10090"
 /clone="NIA:K0976A09 IMAGE:30091496"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated) cDNA Library (Long)"
 /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://gsun.grc.nia.nih.gov/cdna). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Angelo L. Vescovi (Institute for Stem Cell Research, Italy). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pCAGTGTAGTCGAGCGGCCCTTTTITTTTITTTT-3'] from 2.0 Microgram of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.8 kb. The library was constructed by Yulan Piao."

ORIGIN
 Alignment Scores:
 Pred. No.: 1,136-102 Length: 585
 Score: 955.00 Matches: 179
 Percent Similarity: 96.92% Conservative: 10
 Best Local Similarity: 91.79% Mismatches: 6
 Query Match: 48.23% Indels: 0
 DB: 6 Gaps: 0

US-09-983-000a-2 (1-374) x CA879683 (1-585)

Qy 70 LeuThrGlnValAsnValAsnLeuLysLeuLysPheGlnGlyTrpAspLysThrSer 89
 Db 1 CTTACACAAAGTCAATGTGAATCTTAAGAACTGAATTTTCAGGGTTGGGAAAGACGGTCC 60

Qy 90 LeuGluAsnThrPheIleHisAsnThrGlyLysThrValGluIleAsnLeuThrAsnAsp 109
 Db 61 TTGGAAACACACGTTTCATTCACACACAGTGGGAAACACAGTGGAAATATAATCTCACTAATGAC 120
 Qy 110 TyArgValSerGlyLysValSerGluMetValPheLysAlaSerLysIleThrPheHis 129
 Db 121 TACTATCTCAGTGGAGGACTTTTCAGAAAGGCTTTCAGGCGCAAGCAAGATACTTTCAC 180
 Qy 130 TrpGlyLysCysAsnMetSerSerAspGlySerGluHisSerLeuGluGlyGlnLysPhe 149
 Db 181 TGGGAAATGCAATGTGTCATCTGAAGGATCGAACATAGCTTAGAAGGACAGAGTTC 240
 Qy 150 ProLeuGluMetGlnIleTyCysPheAspAlaAsnArgPheSerSerPheGluGluAla 169
 Db 241 CCACCTGAGATGCAAGTCTACTGCTTGTGTCGACAGATTTTCCAGCTTTTGAGGAAGCA 300
 Qy 170 ValLysGlyLysGlyLysLeuArgAlaLeuSerIleLeuPheGluValGlyThrGluGlu 189
 Db 301 GTTAAAGAAAGGAGATTAAAGGCTTATCCATTTTATTTGAGGTTGGAGTTGAAGAA 360
 Qy 190 AsnLeuAspPheLysAlaIleLeuAspGlyValGluSerValSerArgPheGlyLysGln 209
 Db 361 AATTGGAATACAAAGCCATTATTGATGGAACAGAGTGTAGTCGTTTGTGAAAGCAG 420
 Qy 210 AlaAlaLeuAspProPheIleLeuAsnLeuLeuProAsnSerThrAspLysTyTyr 229
 Db 421 GCTGCTTTAGATCCATTCGTTTGCAGAACCTCTCTGCAAACTCCACTGACAGATTATAC 480
 Qy 230 IleTyAsnGlySerLeuThrSerProCysThrAspThrValAspTrpIleValPhe 249
 Db 481 ATTACAAATGATCATTCATGACATCCCTCCCTGCACACACCGTGGATGGATGTTT 540
 Qy 250 LysAspThrValSerIleSerGluSerGlnLeuAlaValPheCys 264
 Db 541 AAGGATACAGTTAGCATCTCTGAAAGCCAGCTGCTGTTATTTGT 585

RESULT 7
 CK781524 718 bp mRNA linear EST 23-FEB-2004
 LOCUS UI-M-GV0-clly-h-20-0-UI.r1 NIH_BMAP_GV0 Mus musculus cDNA clone
 DEFINITION IMAGE:30622171 5', mRNA sequence.
 CK781524
 ACCESSION CK781524.1 GI:42747202
 VERSION EST.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: csapbs@email.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousefl.html
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
 Seq primer: pYX-5.
 Location/Qualifiers
 1..718
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30622171"
 /tissue_type="whole brain"
 /dev_stage="1,5, and 15 days newborn"
 /lab_host="DH10B (T1 phage resistant)"

/clone lib="NIH BMAP_GV0"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CCAATGAT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 2,016-102 Length: 718
Score: 954.00 Matches: 183
Percent Similarity: 94.42% Conservative: 3
Best Local Similarity: 92.89% Mismatches: 11
Query Match: 48.18% Indels: 0
DB: 7 Gaps: 0

US-09-983-000A-2 (1-374) x CK781524 (1-718)

QY 178 AlaLeuSerIleLeuPheGluValGlyThrGluGluAsnLeuAspPheValAlaIle 197
DB 3 GCTTTATCATTTTATTTGAGGTGGAGTTGGAAGAAAATTTGGATTAACAAGCCATTATT 62
QY 198 AspGlyValGluSerValSerArgPheGlyIysGlnAlaLeuAspProPheIleLeu 217
DB 63 CATGGAACGTGAGAGTTAGTCTGTTTGAAGACAGGCTGCTTAGATCCATTGCTTGG 122
QY 218 LeuAsnLeuLeuProAsnSerThrAspIysTrpIleIleIleIleIleIleIleIle 237
DB 123 CAGAACCTCTGCCAACTCCACTGACAAAGTATTACATTTACATGATGATCATTCGATCC 182
QY 238 ProProCysThrAspThrValAspTrpIleValPheIysAspThrValSerIleSerGlu 257
DB 183 CTCTCCTGCACAGACACCGTGAAGTGGATGTTTAAAGGATACAGTTAGCATCTCTGAA 242
QY 258 SerGlnLeuAlaValPheCysGluValLeuThrMetGlnGlnSerGlyTrpValMetLeu 277
DB 243 AGCAGCTGGCTGTATTTGTGAGTTCTCACAATGCAACAGCTCGGTGATGATCATGTTG 302
QY 278 MetAspTrpLeuGlnAsnAsnPheArgGluGlnGlnTrpIleIleIleIleIleIleIle 297
DB 303 ATGGATTACTTACAAAACAATTTCCGAGAACACAGTACAAAGTTTCCAGGCAGGTGTTT 362
QY 298 SerSerTrpThrGlyIysGluIleHisGluAlaValCysSerSerGluProGluAsn 317
DB 363 TCCTCATATCTGGAAGAGAGAGATCCACGAGTAGTGTGTAGTTACAGACCAAGAAAT 422
QY 318 ValGlnAlaAspProGluAsnTrpThrSerLeuLeuValThrTrpGluArgProArgVal 337
DB 423 GTCCAGCTGACCTGAGATTACACAGCTTCTGTCACATGGGAAGACCTCGGGTC 482
QY 338 ValTrpAspThrMetIleGluIysPheAlaValLeuTrpGlnGlnLeuAspGlyGluAsp 357
DB 483 GTTTATGACGCCCATGATTGAGAAGTTTCAGTTCTGTACACGACCATGGCGGGAATGAC 542
QY 358 GlnThrIysHisGluPheLeuThrAspGlyTrpGlnAspLeuValThrIle 374
DB 543 CARGCAAGCATGAGTTCTTAACAGATGCTATCAGGACTTGGGTGCCATT 593

RESULT 8

CA874874

LOCUS CA874874 664 bp mRNA linear EST 20-DEC-2002
DEFINITION K0935E06-SN NIA Mouse Neural Stem Cell (Undifferentiated) cDNA
Library (Long) Mus musculus cDNA clone NIA:K0935E06 IMAGE:30087605

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

5', mRNA sequence.
CA874874
CA874874.1 GI:27326423
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 664)
Piao, Y., Dudekula, D.B., Qian, Y., Martin, P.R., Aiba, K., Vescovi, A.L.
and Ko, M.S.H.
Systematic Analyses of NIA Mouse Neural Stem Cell
(Undifferentiated) cDNA Library (Long)
Unpublished (2002)
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0935 row: 8 column: 06
Seq primer: -21M13 Reverse
High quality sequence stop: 664
POLYA-No.

FEATURES
source

Location/Qualifiers
1..664
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD1"
/db_xref="nia:EST:K0935E06-SN"
/db_xref="taxon:10090"
/clone="NIA:K0935E06 IMAGE:30087605"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated)
cDNA Library (Long)"
/notes="vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
a long-transcript-enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Dr. Angelo L. Vescovi (Institute for Stem
Cell Research, Italy). Double-stranded cDNAs were
synthesized with an Oligo(dT) primer [Invitrogen:
5'-pGATAGTCTAGATCGAGCGCGCCCTTTTCTTTT-3'] from
2.0 Microgram of total RNA, treated with T4 DNA
polymerase, and purified by ethanol-precipitation. The
cDNAs were ligated to lone-linker LL-Sal4, purified by
phenol/chloroform, and separated from free linkers by
Centricon 100. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S. The products were purified
by phenol/chloroform and Centricon 100. The cDNAs were
digested with SalI and NotI enzymes and cloned into
SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.
coli host was transformed with the ligation mixture by the
standard chemical method. The average insert size is about
3.8 kb. The library was constructed by Yulan Piao."

ORIGIN

Alignment Scores:

Pred. No.: 3,076-97 Length: 664
Score: 910.00 Matches: 166
Percent Similarity: 96.70% Conservative: 10
Best Local Similarity: 91.21% Mismatches: 6
Query Match: 45.96% Indels: 0
DB: 6 Gaps: 0

US-09-983-000A-2 (1-374) x CA874874 (1-664)

QY 1 MetArgIleLeuIysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
DB 118 ATGCGAATCTCGAGAGCTTCTCTCGGTGCGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 177

Library (Long) Mus musculus cDNA clone NIA:K0903A03 IMAGE:30084482
5', mRNA sequence.

ACCESSION CA870557
VERSION CA870557.1 GI:27322106
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 522)
Piao, Y., Dudekula, D.B., Qian, Y., Martin, P.R., Aiba, K., Vescovi, A.L.
and K.O.M.S.H.
Systematic Analyses of NIA Mouse Neural Stem Cell
(Undifferentiated) cDNA Library (Long)
Unpublished (2002)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0903 row: A column: 03
Seq primer: -21M13 Reverse
High quality sequence stop: 522
POLYA=No.

FEATURES

source

Location/Qualifiers

1..522
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD1"
/db_xref="taxon:10090"
/clone="NIA:K0903A03 IMAGE:30084482"
/dev_stage="Adult"
/lab_host="PH108"
/clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated)
cDNA Library (Long)"
/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Dr. Angelo L. Vescovi (Institute for Stem
Cell Research, Italy). Double-stranded cDNAs were
synthesized with an oligo(dT) primer [Invitrogen:
5'-pACTAGTCTGATCGGAGCGCGCCCTTTTCTTTT-3'] from
2.0 Microgram of total RNA, treated with T4 DNA
polymerase, and purified by ethanol-precipitation. The
cDNAs were ligated to Lone-linker LL-Sal4, purified by
phenol/chloroform, and separated from free linkers by
Centricon 100. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S. The products were purified
by phenol/chloroform and Centricon 100. The cDNAs were
digested with SalI and NotI enzymes and cloned into
SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.
coli host was transformed with the ligation mixture by the
standard chemical method. The average insert size is about
3.8 Kb. The library was constructed by Yulan Piao."

ORIGIN

Alignment Scores:

Pred. No.: 8 39e-92 Length: 522
Score: 863.00 Matches: 164
Percent Similarity: 97.69% Conservative: 5
Best Local Similarity: 94.80% Mismatches: 4
Query Match: 43.59% Indels: 0
DB: 6 Gaps: 0

US-09-983-000a-2 (1-374) x CA870557 (1-522)

QY

140 SerGluHisSerLeuGluGlyGlnIysPheProLeuGluMetGlnIleTyrCysPheAsp 159

Db 2 TCGGAACATAGCTTAGAAGGACAGAAAGTTCCCACTGGAGATGCAAGTCTACTGCTTTGAT 61
QY 160 AlaSepArgPheSerSerPheGluGluAlaValIysGlyIysGlyIysLeuArgAlaLeu 179
Db 62 GCGGACAGATTTCCAGTTTGGAGAGCAGTTAAAGGAAGGAAGATTAAGGGCTTVA 121
QY 180 SerIleLeuPheGluValGlyThrGluGluAsnLeuAspPheLeysAlaIleIleAspGly 199
Db 122 TCCATTTTATTTGAGTTGGAGTTGAGAAATTTGGATTACAAAGCCATTATTGATGA 181
QY 200 ValGluSerValSerArgPheGlyIysGlnAlaLeuAspProPheIleLeuLeuAsn 219
Db 182 ACTGAGAGTGTAGTCTGTTTGGAAAGCAGCGTCTTTAGATCCATTCTGTCAGAAC 241
QY 220 LeuLeuProAsnSerThrAspIysTyrIleTyrAsnGlySerLeuThrSerProPro 239
Db 242 CTCCTGCCAAATCCACTGACAGTATTACATTTTACATGATGATTCATTCATCCCTCCC 301
QY 240 CysThrAspThrValAspTrrPileValPheIysAspThrValSerIleSerGluSerGln 259
Db 302 TGCACAGACACCGTGGATGATGTTGTTTAAAGGATACAGTTAGCATCTCTGAAAGCCAG 361
QY 260 LeuAlaValPheCysGluValLeuThrMetGlnInSerGlyTyrValMetLeuMetAsp 279
Db 362 CTGGCTGTATTTTGTGAAGTTCTCACATGCAACAGTCTGGGTATGTCATGTTGATGAT 421
QY 280 TyrLeuGlnAsnAsnPheAspGluGluGlnIlyIysPheSerArgGlnValPheSerSer 299
Db 422 TACTTACAAAACAAATTTCCGAGAACACACAGTACAGTTTCCAGCAGGTGTTTTCCTCA 481
QY 300 TyrThrGlyIysGluGluIleHisGluAlaValCysSer 312
Db 482 TATACTGGAAAGGAAGAGATCCACGAGTACTGTGTAGT 520
RESULT 11
CV040479 552 bp mRNA linear EST 01-OCT-2004
LOCUS 438002 BARC 3GAL chicken mixed tissue Gallus gallus cDNA clone
DEFINITION 3GAL17P24 5', mRNA sequence.
ACCESSION CV040479
VERSION CV040479.1 GI:53559918
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 552)
AUTHORS Evocek-Clover, C.M., Ashwell, C.M., McMurtry, J.P., Lillehoj, H.S.,
Matukumalli, L.K. and Van Tassel, C.P.
TITLE Characterization of expressed sequence tags generated from multiple
chicken tissues
JOURNAL Unpublished (2004)
COMMENT Contact: Christina M. Clover
Growth Biology Laboratory
Animal and Natural Resources Institute
Bldg. 200 Rm215 BARC-East, Beltsville, MD 20705, USA
Tel: 3015048224
Fax: 3015048623
Email: chris@anri.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim alt -trim fastavector identified by
cross_match using options -minmatch 12 -minscore 12
Plate: 17 row: F column: 24
Seq primer: CCCAGTCACGACGTGTAAACOG
High quality sequence stop: 552.
FEATURES
Location/Qualifiers
1..552
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Leghorn and broiler"
/db_xref="taxon:9031"
/clone="3GAL_17P24"

```

/lab host="DHSalapha"
/clone_lib="BARC_3GAL chicken mixed tissue"
/note="Vector: pBluescript, SK+, Stratagene; Site_1: NotI;
Site_2: EcoRI; Normalized library of pooled RNA isolated
from whole brain, ultimobronchial gland, parathyroid
gland, cecal tonsil and primordial germ cells Multiple"

ORIGIN
Alignment Scores:
Pred. No.: 4,19e-85 Length: 552
Score: 807.00 Matches: 148
Percent Similarity: 92.90% Conservatve: 22
Best Local Similarity: 80.87% Mismatches: 13
Query Match: 40.76% Indels: 0
DB: 7 Gaps: 0

US-09-983-000A-2 (1-374) x CV040479 (1-552)
Qy 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 4 GAACATAGCTTGGAGAGGACAAAATTTCTCTTGAGATGCAAAATCTACTGCTATGATGGA 63
Qy 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 64 GCCCGCTTTACAGACTTCAAGAGGCAATTAAGGAAATGCAAGTTAAGAGCTTTATCA 123
Qy 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIleAspGlyVal 200
Db 124 GTTTTGTGATTTGGAGTGAAGATAATCCGATTATATTTCCAATCATTAAGGGAGTA 183
Qy 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
Db 184 GATAGTGTGTAGTCGTTTTGGAAAACAGCTGCTTAGAACCATTTGTTTCTGAACCTT 243
Qy 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProCys 240
Db 244 TTACCGATGCAACGACACATAATTACACTTACAATGGTCTTTATCAGCTCCTCCCTGC 303
Qy 241 ThrAspThrValAspTrrIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 304 TCAGAAACAGTTGAATGGATTGTGTTCAAAGATACATAATTAGTATTCTTGAGCAACAGTTA 363
Qy 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 364 GCAGTATTCTGTGAAGTCTTACAAATGCAACAGCTGCTGCTATGATGATCTCATGACATAT 423
Qy 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 424 CTGCAAAACAACTTTTCGAGAGCAACGATATAAGTTCTCTGGGCAAGTGTTTTCTCTTAC 483
Qy 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 484 ACTGGACAGGAAGAAATTCATGAGCAGTTTTCGAGCTCAGAACCTGAAATGTCCAATCT 543
Qy 321 AspProGlu 323
Db 544 GATCCAAAG 552

RESULT 12
CA872941
LOCUS
DEFINITION
K0921G05-5N NIA Mouse Neural Stem Cell (Undifferentiated) cDNA
Library (Long) Mus musculus cDNA clone NIA:K0921G05 IMAGE:30086284
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 498)
AUTHORS
Piao, Y., Dudekula, D.B., Qian, Y., Martin, P.R., Aiba, K., Vescovi, A.L.
```

```

and Ko, M.S.H.
Systematic Analyses of NIA Mouse Neural Stem Cell
(Undifferentiated) cDNA Library (Long)
Unpublished (2002)
COMMENT
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0921 row: G column: 05
Seq primer: -21M13 Reverse
High quality sequence stop: 498
POLYA=No.
FEATURES
Location/Qualifiers
1..498
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD1"
/db_xref="niaEST:K0921G05-5N"
/db_xref="taxon:10090"
/clone="NIA:K0921G05 IMAGE:30086284"
/lab_host="DH10B"
/dev_stage="Adult"
/clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated)
cDNA library (Long)"
/note="vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Dr. Angelo L. Vescovi (Institute for Stem
Cell Research, Italy). Double-stranded cDNAs were
synthesized with an Oligo(dT) primer [Invitrogen:
5'-pGATGTTCTAGATCGGCGCCGCTTTTCTTTT-3'] from
2.0 Microgram of total RNA, treated with T4 DNA
polymerase, and purified by ethanol-precipitation. The
cDNAs were ligated to Lone-linker L1-Sal4, purified by
phenol/chloroform, and separated from free linkers by
Centricon 100. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S. The products were purified
by phenol/chloroform and Centricon 100. The cDNAs were
digested with SalI and NotI enzymes and cloned into
SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.
coli host was transformed with the ligation mixture by the
standard chemical method. The average insert size is about
3.8 kb. The library was constructed by Yulan Piao."
ORIGIN
Alignment Scores:
Pred. No.: 8,17e-85 Length: 498
Score: 804.00 Matches: 151
Percent Similarity: 96.99% Conservatve: 10
Best Local Similarity: 90.96% Mismatches: 5
Query Match: 40.61% Indels: 0
DB: 6 Gaps: 0

US-09-983-000A-2 (1-374) x CA872941 (1-498)
Qy 102 ValGluIleAsnLeuThrAsnAspTyrArgValSerGlyValSerGluMetValPhe 121
Db 1 GTGGAATTAATCTCACTAATGACTATCTCTAGTGAGGACTTTCAGAAAAGTCTTC 60
Qy 122 LysAlaSerLysIleThrPheHisTrrPglyLysCysAsnMetSerSerAspGlySerGlu 141
Db 61 AAGGCAAGCAAGATAACTTTCCACTGGGAAAATGCAATGTGTCTCATCTGAAGGATCGAA 120
Qy 142 HisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAlaAsp 161
Db 121 CATAGCTTAGAAGGACAGAAAGTCCCACTGGAGATGCAAGTCTACTGCTTTGATGGGAC 180
Qy 162 ArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSerIle 181
```

```

181 AGATTTTCCAGTTTTCAGGAAGCAGTTAAAGGAAAGGAAGATTAAAGGCTTTATCCATT 240
182 LeuPheGluValGlyThrGluGluLeuLeuAspPheLeuValAlaIleLeuAspGlyValGlu 201
241 TTATTTGAGTTGGAGTTGAGGAAATTTGGATTACAAAGCCATATTATGGAAGTGA 300
202 SerValSerArgPheGlyLysGlnAlaLaLeuAspProPheLeuLeuLeuLeuLeu 221
301 AGTTTGTAGTCGTTTGGAAAGCAGGCTGCTTTAGATCCATTCGTCGTCGAGAACCTCCTG 360
222 ProAsnSerThrAspLysTyrrIleTyrrAsnGlySerLeuThrSerProProCysThr 241
361 CCAAACTCCACTGACAAGATTATACATTTTACAATGGATCATTCGATCCCTCCCTGCACA 420
242 AspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeuAla 261
421 GACACCGTGATGGATGTTGTTTAAAGGATACAGTTAGCATCTCTGAAGCCAGCTGGCT 480
262 ValPheCysGluValLeu 267
481 GTATTTTGTGAAGTTCTC 498

RESULT 13
BG202130/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 880)
Harrington, J. J., Sherf, B., Rundlett, S., Jackson, P. D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,
Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,
Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K.,
Offenbacher, J., Danzig, J. and Ducar, M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
21227151
11329013
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 384.
FEATURES
source
1..880
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

ORIGIN
Alignment Scores:
Pred. No.: 3 98e-82 Length: 880
Score: 784.50 Matches: 193
Percent Similarity: 82.76% Conservative: 9
Best Local Similarity: 78.88% Mismatches: 38
Query Match: 39.62% Indels: 8

```

```

DB:
US-09-983-000a-2 (1-374) x BG202130 (1-880)
1
4
Gaps:
1
144 LeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrrCysPheAspAlaAspArgPhe 163
859 GTTAAAGCCCAACATTTCCCTTGGGATCGAATTTGTTGTGT---GATGG-GACCGATTT 804
164 SerSerPheGluGluAlaValLysGlyLysGlyLysLeuAtgAlaLeuSerIleLeuPhe 183
803 CT-CAGTTCTGGGACCCAGTCAAGGAAAGAGTAAGTAA--AAGGCTTTACCCATTTGGTTT 746
184 GluValGlyThrGluGluAsnLeuAspPheLysAlaIleLeuAspGlyValGluSerVal 203
745 GAGGTTGGGCCAGAGAAAATTTGAATTTCCAACCAATATTGA--GGAGTCGAAAGTGT 688
204 SerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeuLeuProAsn 223
687 AGTCGTTTGGGAG-CAGCTGGCTTTAAACCAATTTCAAATTTGTTAAACCTTTTCGCGCAAAAC 629
224 SerThrAspLysTyrrIleTyrrAsnGlySerLeuThrSerPro-ProCysThrAspTh 243
628 TCAATTGCCAAGTATTCATTTACATGGCTCATTTGACATCTCTCCCGGCGCAAGACCC 569
243 rValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeuAlaValPh 263
568 AGTTGACTGGATTGT-TTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTGGCTGTTT 510
263 eCysGluValLeuThrMetGlnGlnSerGlyTyrrValMetLeuMetAspTyrrLeuGlnAs 283
509 TTGTGAAGTTCTTCAATGCAACCAATCTGTTATGTCATGCTGATGAGACTACTTACAAA 450
283 nAsnPheArgGluGlnGlnTyrrLysPheSerArgGlnValPheSerSerTyrrThrGly 303
449 CAATTTTCGAGAGCAACAGTACAAAGTTCTCTAGACAGGTGTTTCTCATACACTGTAAA 390
303 eGluGluIleHisGluAlaValCysSerSerGluProGluGlnValGlnAlaAspProGl 323
399 GGAAGAGATTTCATGAAGCAGTTTGTAGTTCAGAACCAAGAAATGTTTCAGGCTGACCCAGA 330
323 uAsnTyrrThrSerLeuLeuValThrTrrGluArgProArgValValTyrrAspThrMetIl 343
329 GAATTTATACAGCCCTCTTGTTCATATGGGAAAGACCTCGAGTCGTTTATGATACCATGAT 270
343 eGluLysPheAlaValLeuTyrrGlnGlnLeuAspGlyLysGlnThrLysHisGluPh 363
269 TGAGAAGTTTCAGTTTGTACACAGCAGTTCGATGGAGGAGGACCAACCAAGCATGAATT 210
363 eLeuThrAspGlyTyrrGlnAspLeuValThrIle 374
209 TTTGACAGATGGCTATCAAGACTTTGGGTGCTATT 176

RESULT 14
CA871074
LOCUS
DEFINITION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Piao, Y., Dudekula, D. B., Qian, Y., Martin, P. R., Alba, K., Vescovi, A. L.
and Ko, M. S. H.
Systematic Analyses of NIA Mouse Neural Stem Cell
(Undifferentiated) cDNA Library (long)
Unpublished (2002)
Contact: Dawood B. Dudekula
Laboratory of Genetics
CA871074
522 bp mRNA linear EST 20-DEC-2002
K0907F04-SN NIA Mouse Neural Stem Cell (Undifferentiated) cDNA
Library (long) Mus musculus cDNA clone NIA:K0907F04 IMAGE:30084927
5', mRNA sequence.
CA871074
CA871074.1 GI:27322623
EST.
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Piao, Y., Dudekula, D. B., Qian, Y., Martin, P. R., Alba, K., Vescovi, A. L.
and Ko, M. S. H.
Systematic Analyses of NIA Mouse Neural Stem Cell
(Undifferentiated) cDNA Library (long)
Unpublished (2002)
Contact: Dawood B. Dudekula
Laboratory of Genetics

```

National Institute on Aging/National Institutes of Health
333 Cassell drive, Suite 400, Baltimore, MD 21224-6820, USA
Email: cdna@gsun.grc.nia.nih.gov
Plate: K0907 row: F Column: 04
Seq primer: -21M13 Reverse
High quality sequence stop: 522
POLYA=No.

FEATURES

1. 522

```
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD1"
/db_xref="niaEST:K0907F04-5N"
/db_xref="taxon:10090"
/clone="NIA:K0907F04 IWAGE:30084927"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated)
cDNA Library (Long)"
/note="vector: pSPOT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Dr. Angelo L. Vescovi (Institute for Stem
Cell Research, Italy). Double-stranded cDNAs were
synthesized with an Oligo(dT) primer [Invitrogen:
5'-pGATGTCCTAGTACGACGCGCGCCCTTTTTTTTTT-3'] from
2.0 microgram of total RNA, treated with T4 DNA
polymerase, and purified by ethanol-precipitation. The
cDNAs were ligated to lone-linker Lu-Sal4, purified by
phenol/chloroform, and separated from free linkers by
Centricon 100. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S. The products were purified
by phenol/chloroform and Centricon 100. The cDNAs were
digested with SalI and NotI enzymes and cloned into
SalI/NotI site of pSPOT1 plasmid vector. The DH10B E.
coli host was transformed with the ligation mixture by the
standard chemical method. The average insert size is about
3.8 kb. The library was constructed by Yulan Piao."
```

ORIGIN

Alignment Scores:		
Pred. No.:	2,768-82	Length:
Score:	783.00	Matches:
Percent Similarity:	96.15%	Conservative:
Best Local Similarity:	90.38%	Mismatches:
Query Match:	35.55%	Indels:
DB:	6	Gaps:
		0
US-09-983-000A-2 (1-374) x CA871074 (1-522)		

US-09-983-000A-2 (1-374) x CA871074 (1-522)

	1	MetArgIleLeuLysAraPheLeuAlaCysIleGlnLeuValCysValCysArgLeuAsp	20
Qy		::: ::: :::	
Db	54	ATGCGAATCTCGACAGACTTCCTCGGTGGCGTCTCAGCTCTGTGTGTGCCTCGCCCTGGAC	113
	21	TtpAlaasnGlyTyrrArgGlnGlnArgLysLeuValGlUGluIleGlyTrpSerTyr	40
Qy		::: ::: :::	
Db	114	TGGGCTTATGGATATTACAGACACAGAGGAAACTTGTTGAAGGATTGGCTGCTCTAC	173
	41	ThrGIyAlaLeuasnGlnLysAsnTrpGlyLysLysTyrProthrCysAsnSerProLys	60
Qy		::: ::: :::	
Db	174	ACAGAGCACCTAAATCAAAAAATTTGGGGAAAGAAAATATCCAATATGTAATAGCCCCAAG	233
	61	GlnSerProIleasnIleaspGluaspleuthrGlnValasnValasnLeuLysLysLeu	80
Qy		::: ::: :::	
Db	234	CAGTCTCCTATTAAATTATGATGAAGATCTTACACAGTCAATGTGAATCTTAAAGAACCTG	293
	81	LysPheGlnGlyTrpAspLysThrSerLeuGluasnThrPhelHisasnThrGlyLys	100
Qy		::: ::: :::	
Db	294	AAATTTTCAGGGTTGGGAAAAACGGCTTTGGAAACACGTTCAATTCACAACACTGGGAAA	353

Qy	101	ThrValGluIleAanLeuThrAenAspTyrArgValSerGlyGlyValSerGluMetVal	120
Db	354	ACAGTGGAAATAATATCTCATATGACTATATCTCAGTGGAGACTTTCAGAAAGGTC	413
Qy	121	PheIysAlaSerIysIleThrPheHisThrPglyLysCysAanMetSerSerAspGlySer	140
Db	414	TTCAAGGCAAGCAAGATAACTTTTCCACTGGGGGAAAATGCAATGTGTCTATGGAAGGATCG	473
Qy	141	GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyr	156
Db	474	GAAACATAGCTTAGAAGGACAGAACTTCCACTGGAGATGCAAGTCTAC	521

RESULT 15

	CD803995	linear	EST 15-JUL-2003
LOCUS	UI-M-GV0-chw-p-13-0-UI.r1 NIH_BMAP_GV0	mRNA	
DEFINITION	Mus musculus cDNA clone IMAGE:30542868 5' mRNA sequence.		

ACCESSION	CD803995	GT:32462821
VERSION	CD803995.1	

KEYWORDS
EST.
Miscellaneous (house mouse)
SOURCE

ORGANISM	Mus musculus
Tissue	Ventricle
Chondrocyte	

Mammalia; Eutheria; Rodent

AUTHORS NIH-MGC <http://mgc.nci.nih>

JOURNAL
Unpublished (1999)

COMMENT
Contact: Robert S. Clausberg
Email: cgarbbs-rc@mail.nih.gov

Tissue Procurement: Dr. JI
CDNA Library Preparation:

CDNA Library Arrayed by:

Clone Distribution: Distr

This clone was contribute

Seq primer: pYX-5.

source
1. .461

```
/organism="Mus mus"
/mol type="mrna"
```

```
/strain="C5/BL/6"
```

```
/clone="IMAGE:305
```

```
/dev_stage="1,5,
```

```
/clone lib="NIH B
```

Site 2: Not I: Th

Bonaguid, Lennon & 1996. Denatured R

gel. First strand
primers containing

size selected acc

directionally int

is CGAACTGAAT. The

Iowa Birth and Developmental Study

INSTITUTE OF MENINGEAL PROGRAM COORDINATOR

ORIGIN

Alignment Scores:

Score: 770.00

Best Local Similarity: 95.39%

```
Query Match: 38.89% Indels: 0
DB: 6 Gaps: 0
US-09-983-000A-2 (1-374) x CD803995 (1-461)

Qy 219 AsnLeuLeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerPro 238
Db 5 AACCTCCTGCCAAACTCCACTGACAAAGTATTACATTTACATGGATCATTGACATCCCT 64

Qy 239 ProCysThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSer 258
Db 65 CCCTGCACAGACCCGTGGATGGATGTTTTAAGGATACAGTTAGCATCTCTGAAAGC 124

Qy 259 GlnLeuAlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMet 278
Db 125 CAGCTGGCTGTATTTTGTGAAGTTCTCAAAATGCACAGTCTGGGTATGTCATGTTGATG 184

Qy 279 AspTyrLeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSer 298
Db 185 GATTACTTTACAAAACAATTTCCGAGAACAAACAGTACAAAGTTTTCCAGGCAGGTGTTTCC 244

Qy 299 SerTyrThrGlyLysGluLulileHisGluAlaValCysSerSerGluProGluAsnVal 318
Db 245 TCATATACTGGAAAGAGAGATCCACGAAAGTAGTGTGTAGTTCAGAACCCAGAAAATGTG 304

Qy 319 GlnAlaAspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValVal 338
Db 305 CAAGCTGACCCCTGAGAAATTACACAGCCTTCTGGTTCACATGGGAAAGACCTCGGTCCTT 364

Qy 339 TyrAspThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGln 358
Db 365 TATGACGCCCATGATTGAGAAGTTTCAGTTCTGTACCAAGCCACTGGCGGAAATGACCAA 424

Qy 359 ThrLysHisGluPheLeuThrAspGlyTyrGlnAsp 370
Db 425 GCCAAGCATGAGTCTTTAACAGATGGCTATCAGGAC 460
```

Search completed: August 30, 2005, 12:01:47
Job time : 3095 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 30, 2005, 06:16:45 ; Search time 560 Seconds

(without alignments)
3953.542 Million cell updates/sec

Title: US-09-983-000A-2

Perfect score: 1980

Sequence: 1 MRILKRFLACIQLLCVCRLD.....GEOQTKHEFLTDGYQDLVTI 374

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of Hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/Cgn2_1/USPTO_spool_p/US09983000/runat_29082005_155923_14661/app_query.fasta_1.519
-DB=N_Geneseq_16Dec04 -QWTF=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human4.0.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09983000 @CGN 1.1.708 @runat_29082005_155923_14661 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004bs: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1980	100.0	3091	8 AAD48126	Aad48126 Human pro
2	1968	99.4	3340	11 ADN39729	Adn39729 Cancer/an
3	1968	99.4	4220	11 ADN39731	Adn39731 Cancer/an
4	1968	99.4	4347	11 ADN39724	Adn39724 Cancer/an
5	1968	99.4	4480	11 ADN39730	Adn39730 Cancer/an

6	1968	99.4	5343	8 ABX76418	Abx76418 Lung canc
7	1968	99.4	5343	8 ABX76219	Abx76219 Lung canc
8	1968	99.4	5343	11 ADN39057	Adn39057 Cancer/an
9	1968	99.4	5367	8 ABX76417	Abx76417 Lung canc
10	1968	99.4	5367	8 ABX76218	Abx76218 Lung canc
11	1968	99.4	5367	11 ADN39055	Adn39055 Cancer/an
12	1968	99.4	5367	11 ADN39726	Adn39726 Cancer/an
13	1968	99.4	5481	8 ABX76420	Abx76420 Lung canc
14	1968	99.4	5481	8 ABX76221	Abx76221 Lung canc
15	1968	99.4	5481	11 ADN39061	Adn39061 Cancer/an
16	1968	99.4	6924	2 AAQ67283	Aaq67283 Human RPT
17	1968	99.4	7941	6 ABL66502	Ab166502 Lung canc
18	1968	99.4	7941	8 ABX76421	Abx76421 Lung canc
19	1968	99.4	7941	8 AAD48125	Aad48125 Human PTP
20	1968	99.4	7941	10 ACC72641	Acc72641 Human pro
21	1968	99.4	7941	11 ADN39063	Adn39063 Cancer/an
22	1968	99.4	7941	12 ADJ75139	Adj75139 Marker ge
23	1968	99.4	7944	8 ABX76217	Abx76217 Lung canc
24	1968	99.4	7944	8 ABX76416	Abx76416 Lung canc
25	1968	99.4	7944	11 ADN39725	Adn39725 Cancer/an
26	1968	99.4	7944	11 ADN39053	Adn39053 Cancer/an
27	1968	99.4	8058	8 AAD48127	Aad48127 Human pro
28	1968	99.4	8292	12 ADQ22702	Adq22702 Human sof
29	1962	99.1	8146	4 AAH73240	Aah73240 Human cer
30	1959	98.9	3350	12 ADQ18133	Adq18133 Human sof
31	1955	98.7	5481	11 ADN39728	Adn39728 Cancer/an
32	1940	98.0	5363	8 ABX76220	Abx76220 Lung canc
33	1940	98.0	5363	8 ABX76419	Abx76419 Lung canc
34	1940	98.0	5363	11 ADN39059	Adn39059 Cancer/an
35	1940	98.0	5363	11 ADN39727	Adn39727 Cancer/an
36	1857	93.8	6801	10 ADB79880	Adb79880 Rat phosp
37	1857	93.8	6801	12 ADO31207	Ado31207 Rat phosp
38	1831	92.5	6887	12 ADJ76009	Adj76009 Marker ge
39	1831	92.5	6887	12 ADO31209	Ado31209 Mouse pho
40	648.5	32.8	4338	2 AAQ71842	Aaq71842 Human RPT
41	648.5	32.8	6687	12 ADQ22975	Adq22975 Human sof
42	412	20.8	813	2 AAT15450	Aat15450 Lung canc
43	412	20.8	822	2 AAT15451	Aat15451 Lung canc
44	412	20.8	986	2 AAT15454	Aat15454 HCAVIII p
45	412	20.8	986	2 AAT15449	Aat15449 Lung canc

ALIGNMENTS

RESULT 1
AAD48126
ID AAD48126 standard; DNA; 3091 BP.

XX AAD48126;

XX AC AAD48126;

DT 24-FEB-2003 (first entry)

XX Human protein tyrosine phosphatase zeta SM1 variant DNA.

Brain; tumour protein target; Tbt; ischaemic stroke; cancer; epilepsy;
schizophrenia; depression; Alzheimer's disease; Parkinson's disease;
Huntington's chorea; traumatic head injury; dementia; stupor; headache;
coma; vertigo; weakness; myasthenia gravis; cerebrovascular disorder;
infection; multiple sclerosis; pregnancy; medical illness; vasotropic;
metabolic deficiency; cerebroprotective; antidepressant; antibacterial;
cytostatic; nootropic; analgesic; fungicide; virucide; human; enzyme;
protein tyrosine phosphatase zeta SM1; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 148..1272

FT /*tag= a

FT /product= "Human protein tyrosine phosphatase zeta SM1

FT variant"

FT misc_feature 1262

FT /*tag= b

FT /note= "Coding sequence alternative splice site"

FT 3'UTR 1273..3091
 XX /*tag= c
 PN WO200276510-A1.
 XX
 XX
 XX
 XX PD 03-OCT-2002.
 XX
 XX PF 22-MAR-2002; 2002WO-US008992.
 XX
 XX PR 23-MAR-2001; 2001US-00816703.
 XX PR 17-OCT-2001; 2001US-00983000.
 XX
 XX PA (AGVT-) AGY THERAPEUTICS INC.
 XX
 XX
 XX PI Mueller S, Melcher T, Chin DJ;
 XX P-PSDB; AAE30333.
 DR WPI; 2003-029903/02.
 XX
 XX PT Developing active agents that modulate the activity of a brain tumor
 PT protein target gene or gene product for treating e.g. stroke or cancer,
 PT comprises contacting an agent with a brain tumor protein.
 XX
 XX PS Claim 1; Page 81-84; 135pp; English.
 XX
 CC The invention relates to a method for developing biologically active
 CC agents that modulate activity of a brain tumour protein target (Tbt) gene
 CC or gene product. The method is useful for developing biologically active
 CC agents that modulate the activity of a brain tumour protein target gene
 CC or gene product. Compounds that bind to the brain tumour proteins are
 CC useful for treating e.g. ischaemic stroke, brain cancer, epilepsy,
 CC schizophrenia, depression, Alzheimer's disease, Parkinson's disease,
 CC Huntington's chorea, traumatic head injury, dementia, stupor, headache,
 CC coma, vertigo, weakness, myasthenia gravis, cerebrovascular disorders,
 CC infectious disorders (including fungal, bacterial, viral and parasitic
 CC infections), multiple sclerosis, and other complications associated with
 CC pregnancy, medical illness, alcohol and substance abuse, toxins and
 CC metabolic deficiencies. The brain tumour proteins may also be used to
 CC raise antibodies. The present sequence is human protein tyrosine
 CC phosphatase zeta SMI variant DNA used to illustrate the method of the
 CC invention
 XX
 SQ Sequence 3091 BP; 1008 A; 535 C; 590 G; 958 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.52e-207 Length: 3091
 Score: 1980.00 Matches: 374
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0
 US-09-983-000A-2 (1-374) x AAD48126 (1-3091)
 QY 1 MetArgIleLeuLysArgPheLeuAlaCyatIleGlnLeuLeuCysValCysArgLeuAsp 20
 DB 148 ATGCGAATCCTTAAGCGGTTCTCTCGCTTGCATTCAGCTCCTCTGTGTGGCGCTGGAT 207
 QY 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
 DB 208 TGGGCTAATGGATACTACAGACACACAGAGAAACCTTGTGTAAGAGATTGGCTGCTCTAT 267
 QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
 DB 268 ACAGGAGCACTGAATCAAAATAAATGGGGAAGAAATATCAACATGTAATAGCCCAAAA 327
 QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
 DB 328 CAATCTCTATCAATATTGTAGAGATCTTACACAAAGTAATGTGAATCTTAAGAAACTT 387
 QY 81 LysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
 DB 388 AAATTTGAGGGTTGGATAAAACATCATTTGGAATAACACATTCATTCATAAACACTGGGAA 447

QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
 DB 448 ACAGTGGAAATTAATCTCTCACTAATGACTACCGTGTGTCAGCGGAGGATTTTCAGAAATGGTG 507
 QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
 DB 508 TTTAAAGCAAGCAGATAACTTTTCACTGGGAAATGCAATATGTCTCATCTGATGGATCA 567
 QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
 DB 568 GAGCATAGTTTAGAAGGACAAAATTTCCACTTGAGATGCAAAATCTACTTGTCTTGTATGCG 627
 QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
 DB 628 GACCGATTTTCAAGTTTGGAGGACAGCTCAAGGAAAGGAGTAAAGAGCTTTATATCC 687
 QY 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleLeuAspGlyVal 200
 DB 688 ATTTTGTGTTGAGGTTGGACAGAGAAATTTGGATTTCAAGCGCATATTGATGGAGTC 747
 QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspPropheIleLeuLeuAsnLeu 220
 DB 748 GAAAGTGTAGTCTGTTTGGAGAGCAGGCTGCTTTAGATCCATTCATCTACTGTGTAACCTT 807
 QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240
 DB 808 CTGCCAAACTCAACTGACAAAGTATTACATTTACAAATGGCTCATTTGACATCTCTCCCTCC 867
 QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
 DB 868 ACAGACACAGTTGACTGCTGATGTTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
 QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
 DB 928 GCTGTTTTTGTGAAGTTCTTACAAATGCAACAATCTGTTATGTATGTCATGCTGATGACTAC 987
 QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
 DB 988 TTACAAACAATTTTCGAGAGCAACAGTCAAGTTCTCTAGACAGGTGTTTTCTCTCATAC 1047
 QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
 DB 1048 ACTGGAAGGAGAGAGATTTCATGAGCAGTTGTGTAGTTTCAGAACACAGAAATGTCAGCT 1107
 QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
 DB 1108 GACCCAGAGAATTATACAGCCTTCTTGTGTACATGGGAAAGACCTCGAGTGTGTTATGAT 1167
 QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
 DB 1168 ACCATGATTGAGAAAGTTTGCAAGTTTTGTACCAAGTTGATGGAGAGGACCAAAACCAAG 1227
 QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
 DB 1228 CATGAATTTTGACAGATGGCTATCAAGACTTGGTAACTATA 1269
 RESULT 2
 ADN39729
 ID ADN39729 standard; cDNA; 3340 BP.
 XX
 AC ADN39729;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:CI01.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;

XX vulnerable; gene therapy; vaccine; gene; ss.

XX Homo sapiens.

XX WO2003042661-A2.

XX PD 22-MAY-2003.

XX PF 13-NOV-2002; 2002WO-US036810.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 21-NOV-2001; 2001US-0332464P.

XX PR 29-NOV-2001; 2001US-0334393P.

XX PR 03-DEC-2001; 2001US-0335394P.

XX PR 14-DEC-2001; 2001US-0340376P.

XX PR 08-JAN-2002; 2002US-0347211P.

XX PR 10-JAN-2002; 2002US-0347349P.

XX PR 08-FEB-2002; 2002US-0352509P.

XX PR 13-FEB-2002; 2002US-0356714P.

XX PR 20-FEB-2002; 2002US-0359077P.

XX PR 29-MAR-2002; 2002US-0368809P.

XX PR 04-APR-2002; 2002US-0370110P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PR 05-JUN-2002; 2002US-0386614P.

XX PR 16-JUL-2002; 2002US-0396839P.

XX PR 22-JUL-2002; 2002US-0397759P.

XX PR 22-JUL-2002; 2002US-0397845P.

XX PR 09-SEP-2002; 2002US-0409450P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Revezzi PA;

XX PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX DR WPI: 2003-468649/44.

XX DR P-PSDB; ADN39946.

XX Determining the presence or absence of a pathological cell in a patient, useful for diagnosing, prognosing or treating cancer, comprises detecting a nucleic acid in a biological sample.

XX Claim 8; SEQ ID NO C101; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064) whose expression is upregulated or downregulated in specific cancers or other diseases such as angiogenic or fibrotic disorders, and to methods of determining the presence or absence of a pathological cell in a patient by detecting a nucleic acid at least 80% identical to those of the invention or by detecting a polypeptide of the invention. The invention also relates to expression vectors and host cells comprising a nucleic acid of the invention; antibodies which specifically bind a polypeptide of the invention; use of such antibodies for drug targeting; and methods of screening for modulators of activity or expression of the polypeptides and nucleic acids. The nucleic acids, polypeptides, antibodies and methods are useful for diagnosing, prognosing and treating cancer and other conditions such as psoriasis, ischaemia, heart disease, atherosclerosis, inflammatory diseases, autoimmune diseases, retinal neovascularisation syndrome, scarring and uterine fibroids. They may also be useful in wound healing and in contraception. The present sequence represents a nucleic acid sequence of the invention.

XX SQ Sequence 3340 BP; 1050 A; 651 C; 746 G; 893 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.:	3,6e-206	Length:	3340
Score:	1968.00	Matches:	372
Percent Similarity:	99.47%	Conservative:	0
Best Local Similarity:	99.47%	Mismatches:	2
Query Match:	99.39%	Indels:	0
DB:	11	Gaps:	0

XX US-09-983-000A-2 (1-374) x ADN39729 (1-3340)

QY	1	MetArgIleLeuLysAArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp	20
DB	1	ATGCGAATCCTAAAGGGTTTCTCGCTTGCATTACAGTCTCTCTGTGTTGGCCCTGGAT	60
QY	21	TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr	40
DB	61	TGGGCTAATGGATACACAGACAACAGAGAAACACTTGTGTGAAGAGATTGGTCTCTAT	120
QY	41	ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys	60
DB	121	ACAGGAGCACTGAATCAAAAATTTGGGGAAGAAATATCCAAATGTAATAGCCCAAAA	180
QY	61	GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu	80
DB	181	CAATCTCTATCAATATTGATGAAGATCTTACACAAGTAAATGTGAATCTTAAGAACTT	240
QY	81	LysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys	100
DB	241	AAATTTTCAGGGTGGGATAAACATCATTTGGGAAACACATTCATTATCAACACTGGGAA	300
QY	101	ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal	120
DB	301	ACAGTGGAAATTAATCTCAATGACATCCGTGTTCAGCGAGGAGTTTCAGAAATGGTG	360
QY	121	PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer	140
DB	361	TTTAAAGCAAGCAAGATAACTTTTTCACGTGGGAAATTCATATGTCTATCTGATGATCA	420
QY	141	GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla	160
DB	421	GAGCATAGTTTAGAAGGACAAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATCG	480
QY	161	AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer	180
DB	481	GACCGATTTTCAAGTTTGGAGAGCAGTCAAGGAAAGGAGTTTAAGAGCTTTATGCC	540
QY	181	IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIleAspGlyVal	200
DB	541	ATTTTGTGAGTTGGGACAGAGAAATTTGGATTTCAAGCGATTTATTGATGAGTTC	600
QY	201	GluSerValSerArgPheGlyLysGlnAlaLeuAspPropheIleLeuLeuAsnLeu	220
DB	601	GAAGTGTGTAGTCTGTTTGGGAGCAGCTCTTTAGATCCATTCATCATCTGTGAACCTT	660
QY	221	LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys	240
DB	661	CTGCCAACTCACTGCAAGTATTACATTTACAAATGCTGTTATGTCTGATGACTAC	720
QY	241	ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu	260
DB	721	ACAGACACAGTTGACTGGATTGTTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG	780
QY	261	AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr	280
DB	781	GCTGTTTTTGTGAAGTTCTTACAAATGCAAACTGCTGTTATGTCTGATGACTAC	840
QY	281	LeuGlnAsnAsnPheArgGluGlnGlyTyrLysPheSerArgGlnValPheSerSerTyr	300
DB	841	TTACAAACAAATTTTCGAGAGCAACAGTACAAAGTTCTCTAGACAGGTGTTTTCTCATAC	900
QY	301	ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla	320
DB	901	ACTGGAAGGAGAGATTCATGAAGCAGTTGTGATGTTTCAGAACCCAGAAATGTCAGGCT	960
QY	321	AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp	340
DB	961	GACCCAGAGAAATTAACAGCTTCTTGTGTACATGGGAAAGACCTCGAGTCTGTTATGAT	1020
QY	341	ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys	360
DB	1021	ACCATGATTGAGAAAGTTGACAGTTTGTACACAGTGTGGATGGAGAGACCAACCAAG	1080
QY	361	HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle	374

||||| 1081 CATGAATTTTGACAGATGGCTATCAAGACTTGGTGCTATT 1122
RESULT 3
ADN39731
ID ADN39731 standard; cDNA; 4220 BP.
XX
AC ADN39731;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:CI03.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerable; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0352509P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368099P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397779P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Azi¹ N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DR, Mutray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR WPI; 2003-468649/44.
DR P-PSDB; ADN39948.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 8; SEQ ID NO CI03; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,

CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
SQ Sequence 4220 BP; 1329 A; 873 C; 912 G; 1106 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.07e-206 Length: 4220
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 11 Gaps: 0

US-09-983-000A-2 (1-374) x ADN39731 (1-4220)
Qy 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
Db 1 ATGCGAATCCTAAAGCGTTTCTCGCTTGCATTGCACTCTCTGTGTGTTCGCGCTCGAT 60
Qy 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40
Db 61 TGGCTAATGGATCTACTACAGACAACAGAAACTTGTGAGAGATTGGCTGGTCTTAT 120
Qy 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysTyrProThrCysAsnSerProLys 60
Db 121 ACAGGAGCACTGAATCAAAAAAATTGGGAAAGAAATATCAACATGTAATAGCCCAAAA 180
Qy 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLeu 80
Db 181 CAATCTCTATCAATATTTGATGAAGATCTTACCAAGTAATGTAATCTTAAGAAACTT 240
Qy 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db 241 AATTTGAGGGTGGGATTAACAATCATTTGGAAAAACACATCTTCAATACACTGGGAAA 300
Qy 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
Db 301 ACAGTGGAAATTAATCTCACTAATGACTACCGTGTGACGCGAGGAGTTTCAGAAATGGT 360
Qy 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
Db 361 TTTAAAGCAAGCAAGATAACTTTTCACTGGGAAAAATGCAATATGTCTCATCTGATGGATCA 420
Qy 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 421 GAGCATAGTTTAGNAGGACAAAATTTCCACTTGAGATGCAATCTACTCTTGTGATGG 480
Qy 161 AspArgPheSerSerPheGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 481 GACCGATTTTCAAGTTTTTGAGGAAGCAGTCAAGAGAAAAGGAAGTTTAAGAGCTTTATCC 540
Qy 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleAspGlyVal 200
Db 541 ATTTTGTGTGAGGTGGGACAGAAAATTTGGATTTCAAGCGATTTATTGATGGAGTC 600
Qy 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
Db 601 GAAAGTGTAGTCGTTTTGGGAGCAGCGCTGCTTTAGATCCATTCATCTGTTGAACCTT 660
Qy 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240
Db 661 CTGCCAAACTCAACTGACCAAGTATTACATTTACAAATGGCTCATTTGACATCTCTCCCTGC 720
Qy 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 721 ACAGACACAGTTGACTGGATGTTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 780
Qy 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280

Db 781 GCTGTTTTTGTGAAGTTCTTACAATGCAACAACTCTGGTTATGTCTATGCTGATGACTAC 840
 Qy 281 LeuGlnAsnAspPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
 Db 841 TTACAAACAACTTTTCGAGGACACAGTACAAAGTTCTCTAGACAGGTGTTTCTCTATAC 900
 Qy 301 ThrGlyLysGluGluLeuHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
 Db 901 ACTGGAAGGAGAGATTATCATGACGAGTTTGTAGTTCAGAACCAAGAAATGTTCCAGCT 960
 Qy 321 AspProGluAsnTyrThrSerLeuLeuValThrTyrGluArgProArgValValTyrAsp 340
 Db 961 GACCCAGAGAAATATACCAAGCTTCTTGTATCATCGGAAGACCTCGAGTCGTTTATGAT 1020
 Qy 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
 Db 1021 ACCATGATGAGAGTTTGCAGTTTGTACCAAGCAGTTGGATGGAGAGACCAACCAAG 1080
 Qy 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
 Db 1081 CATGAATTTTGACAGATGGCTATCAAGACTTGGGTGCTATT 1122

RESULT 4
 ID ADN39724 standard; cDNA; 4347 BP.
 XX AC ADN39724;
 XX DT 17-JUN-2004 (first entry)
 XX DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO: C96.
 XX KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.
 XX OS Homo sapiens.
 XX FN WO2003042661-A2.
 XX PD 22-MAY-2003.
 XX PF 13-NOV-2002; 2002WO-US036810.
 XX PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-FEB-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-039775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KB, Zlotnik A;
 XX WPI; 2003-468649/44.

P-PSDB; ADN39941.
 Determining the presence or absence of a pathological cell in a patient,
 useful for diagnosing, prognosing or treating cancer, comprises detecting
 a nucleic acid in a biological sample.
 Claim 8; SEQ ID NO C96; 1385pp; English.
 The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 whose expression is upregulated or downregulated in specific cancers or
 other diseases such as angiogenic or fibrotic disorders, and to methods
 of determining the presence or absence of a pathological cell in a
 patient by detecting a nucleic acid at least 80% identical to those of
 the invention or by detecting a polypeptide of the invention. The
 invention also relates to expression vectors and host cells comprising a
 nucleic acid of the invention; antibodies which specifically bind a
 polypeptide of the invention; use of such antibodies for drug targeting;
 and methods of screening for modulators of activity or expression of the
 polypeptides and nucleic acids. The nucleic acids, polypeptides,
 antibodies and methods are useful for diagnosing, prognosing and treating
 cancer and other conditions such as psoriasis, ischaemia, heart disease,
 atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 neovascularisation syndromes, scarring and uterine fibroids. They may
 also be useful in wound healing and in contraception. The present
 sequence represents a nucleic acid sequence of the invention.
 Sequence 4347 BP; 1371 A; 891 C; 939 G; 1146 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. NO.: 5.29e-206 Length: 4347
 Score: 1968.00 Matches: 372
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 2
 Query Match: 99.39% Indels: 0
 DB: 11 Gaps: 0
 US-09-983-000A-2 (1-374) x ADN39724 (1-4347)
 Qy 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
 Db 1 ATGCGAATCTTAAAGCGTTTCTCGCTTGCATTGAGTCTCTCTGTTTGGCGCTGGAT 60
 Qy 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40
 Db 61 TGGCTTAATGGATACTACACACACAGAAAACCTTGTGAGAGATGGTGGTCTCTAT 120
 Qy 41 ThrGlyAlaLeuAsnGlnLysAsnTyrGlyLysLysTyrProThrCysAsnSerProLys 60
 Db 121 ACAGGACACTGAATCAAAAAAATGGGAAAGAAATATCCACATGTAATAGCCCAAA 180
 Qy 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
 Db 181 CAATCTCTATCAATATTTGATGAAGATCTTACACAGTAATGGAATCTTAAAGAACTT 240
 Qy 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
 Db 241 AAATTCAGGGTGGGATAAACATCATTTGGAAAACACATTCATTCATAACACTGGGAAA 300
 Qy 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
 Db 301 ACAGTGGAAAATTAATCTCTCACTAATGACTACCGTGTGACGCGAGGAGTTTCAGAAAATGTG 360
 Qy 121 PheLysAlaSerLysIleThrPheHisTyrGlyLysCysAsnMetSerSerAspGlySer 140
 Db 361 TTTAAAGCAAGCAAGATAACTTTTCACTGGGAAAATGCAATATGTCTATCTGATGGATCA 420
 Qy 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
 Db 421 GAGCATAGTTTAGAGGACAAAATTTCCATCTGAGATGCAATCTACTCTGTTGATGCA 480
 Qy 161 AspArgPheSerSerPheGluAlaValLysGlyLysGlyLysLeuAlaLeuSer 180
 Db 481 GACCGATTTTCAAGTTTTTGAGGAAGCAGTCAAGGAAAAGGAAGTTAAGAGCTTTATTC 540

```
QY 191 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLeuAlaIleAspGlyVal 200
Db 541 ATTTGTTTGGAGTTGGACAGAGAAAATTTGGATTTCAAGGCGATTTATTGATGGAGTC 600
QY 201 GluSerValSerArgPheGlyGlyGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
Db 601 GAAAGTGTTAGTCGTTTGGAGAGCGGCTGCTTTAGATCCATTTCACTACTGTTGAACCTT 660
QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240
Db 661 CTGCCAAACTCAACTGACAGAGTATTACATTTTACAATGCGTCATTTGACATCTCTCCCTGC 720
QY 241 ThrAspThrValAspTyrIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 721 ACAGACACAGTTGACTGGATTTGTTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 780
QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 781 GCTGTTTTTTGTGAAGTTCTTACAATGCAACAATCTGTTATGATGCTGATGACTAC 840
QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 841 TTACAAAACAATTTTCGAGAGCAACAGTACAAAGTTCTCTGACAGGTTGTTTCTCTCATC 900
QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 901 ACTGGAAGAGGAGAGATTCATGAGCAGTTTGTAGTTCCAGAACCCAGAAAATGTTCAGGCT 960
QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTyrGluArgProArgValValTyrAsp 340
Db 961 GACCCAGAGAATTTATCCAGGCTTCTGTTATGATGGGAAGACCTCGAGTCTGTTATGAT 1020
QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 1021 ACCATGATTGAGAGTTTGCAGTTTGTACCCAGCAGTTTGGATGGAGAGACCAACCAAG 1080
QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db 1081 CATCAATTTTTCACAGATGGCTATCAAGACTTGGGTGCTATT 1122
RESULT 5
ADN39730
ID ADN39730 standard; cDNA; 4480 BP.
XX
AC ADN39730;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO: C102.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
FN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
```

```
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
FA (BOSB-) BOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Gingsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR WPI: 2003-468649/44.
DR P-PSDB; ADN39947.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 8; SEQ ID NO C102; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides, and
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
SQ Sequence 4480 BP; 1399 A; 936 C; 968 G; 1177 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 5,53e-206 Length: 4480
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 11 Gaps: 0
```

US-09-983-000A-2 (1-374) x ADN39730 (1-4480)

```
Qy 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
Db 1 ATGCGAATCTTAAAGCGTTTCTCGCTTGCATTGAGTCTCTGTTTGGCGCTGGAT 60
Qy 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40
Db 61 TGGGCTTAATGGATCTACTACAGACACAGAGAAAATCTGTTGAAGAGATGGCTGCTCTAT 120
Qy 41 ThrGlyAlaLeuAsnGlnLysAsnTyrGlyLysLysTyrProThrCysAsnSerProLys 60
Db 121 ACAGAGCAGCTGATCAAAAAAATTTGGGAAAGAAATATCCACATGTAATAGCCCAAA 180
Qy 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLeu 80
Db 181 CAATCTCTATCAATATTGATGAAGATCTTACACAGTAAATGTAATCTTAAAGAACTT 240
```

QY 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAenThrPheIleHisAenThrGlyLys 100
 Db 241 AAATTTCCGGGTGGGATAAAACATCATTTGGAAACACATTCATTATAACACTGGGAAA 300
 QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyValSerGluMetVal 120
 Db 301 ACAGTGGAAATTAATCTCATTAATCACTACCGTGTGACGGAGGAGTTTCAGAAATGGTG 360
 QY 121 PheLysAlaSerLysIleThrPheHisTyrGlyLysCysAsnMetSerSerAspGlySer 140
 Db 361 TTTAAAGCAAGCAAGATAACTTTTCACTGGGAAATGCATATGTCATCTGATGATCA 420
 QY 141 GluHisSerLeuGluGlyClnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
 Db 421 GAGCATAGTTTGAAGAGGCAAAATTTCCACTTGAGATGCAATCTACTGCTTTGATGCA 480
 QY 161 AspArgPheSerSerPheGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
 Db 481 GACCCGATTTTCAAGTGTGAGGAAGCAGTCAAGGAAAGGGAAGTTAAGAGCTTATACC 540
 QY 181 IleLeuPheGluValGlyThrGluGluAenLeuAspPheLysAlaIleIleAspGlyVal 200
 Db 541 ATTTTGTGAGTTGGGACAGAGAAATTTGGATTTCAAGCGATTTATGATGGAGTC 600
 QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
 Db 601 GAAAGTGTAGTGTGTTGGGAGCAGGCTGCTTAGATCCATTCATCATCTGTTGAACCTT 660
 QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProCys 240
 Db 661 CTGCCAATCTCAACTGACAGATTTACATTTACATGTTTCAATGATCTCTCCCTGTC 720
 QY 241 ThrAspThrValAspTyrIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
 Db 721 ACAGACACAGTTGACTGATGTTTAAAGATACAGTTAGCACTCTGAAAGCCAGTTG 780
 QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
 Db 781 GCTGTTTTTGTGAAGTTCTTACATGCAACAATCTGGTTATGTCATGCTGATGACTAC 840
 QY 281 LeuGlnAenAsnPheArgGluGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
 Db 841 TTACAAAAACAATTTTCGAGAGCAACAGTACAGTTCTCTAGACAGGTTTCTCTCATAC 900
 QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
 Db 901 ACTGGAAAGGAGAGATTCATGACAGCTTTGTAGTTCAGAACCAAGAAATGTCAGGCT 960
 QY 321 AspProGluAenTyrThrSerLeuLeuValThrTyrGluArgProArgValValTyrAsp 340
 Db 961 GACCCAGAGAAATTAACACAGCTTCTGTTATGATGGGAAAGACCTTCGAGTCTGTTATGAT 1020
 QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
 Db 1021 ACCATGATTTGAAGTTTTCAGATTTTGTACAGCAGTTTGGATGGAGAGACCAACCAAG 1080
 QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
 Db 1081 CATGAATTTTTCAGAGATGGCTATCAAGACTTTGGGTGCTATT 1122

RESULT 6

ABX76418

ID ABX76418 standard; DNA; 5343 BP.

XX ABX76418;

AC ABX76418;

XX 02-APR-2003 (first entry)

XX Lung cancer-associated polynucleotide #282.

DE Lung cancer-associated polynucleotide #282.

XX Lung cancer-associated polynucleotide; gene; db; cytostatic; emphysema;

KW anti-inflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;

KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;

chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

Unidentified.

WO200286443-A2.

31-OCT-2002.

18-APR-2002; 2002WO-US012476.

18-APR-2001; 2001US-0284770P.

10-MAY-2001; 2001US-0290492P.

09-NOV-2001; 2001US-0339245P.

13-NOV-2001; 2001US-0350666P.

29-NOV-2001; 2001US-0334370P.

12-APR-2002; 2002US-0372246P.

(EOSB-) EOS BIOTECHNOLOGY INC.

Aziz N, Murray R;

WPI; 2003-093161/08.

P-PSDB; ABU56689.

Detecting a lung cancer-associated transcript in a cell from a patient

for treating lung cancer, by contacting a biological sample from the

patient with a polynucleotide that exhibits increased or decreased

expression in lung cancer.

Claim 22; Page 406-407; 453pp; English.

The invention relates to a method for detecting a lung cancer-associated

transcript in a cell from a patient, comprising contacting a biological

sample from the patient with a polynucleotide that selectively hybridizes

to a sequence that is at least 80 % identical to a gene that exhibits

increased or decreased expression in lung cancer samples. Lung cancer-

associated polynucleotides and polypeptides are used for identifying a

compound that modulates a lung cancer-associated polypeptide, for

inhibiting proliferation of a lung cancer-associated cell to treat lung

cancer in a patient and for treating a mammal having lung cancer by

administering a modulatory compound identified. The methods are useful

for treating lung cancer, such as small cell lung cancer, non-small cell

lung cancer or other benign or precancerous lesions, e.g. atelectasis,

emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,

hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and

bronchiectasis. The genes, polynucleotides and polypeptides are useful

for diagnostic purposes and as targets for screening for therapeutic

compounds that modulate lung cancer, such as antibodies. Sequences

CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the

invention

XX Sequence 5343 BP; 1706 A; 1058 C; 1092 G; 1487 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7,17e-206 Length: 5343

Score: 1968.00 Matches: 372

Percent Similarity: 99.47% Conservativeness: 0

Best Local Similarity: 99.47% Mismatches: 2

Query Match: 99.39% Indels: 0

DB: 8 Gaps: 0

US-09-983-000A-2 (1-374) x ABX76418 (1-5343)

Qy 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20

Db 148 ATGCGAATCTTAAAGCGTTTCTCGCTTGCATTCAGCTCTCTGTTTGGCGCTGGAT 207

Qy 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40

Db 208 TGGGCTAATGGATCTACTACGACACAGAGAAACTTGTGAGAGAGATGGCTGCTCAT 267

Qy 41 ThrGlyAlaLeuAenGlnLysAsnTyrGlyLysTyrProThrCysAsnSerProLys 60

Db 148 ATCCGAATCTCTAAAGCGTTTCCTCGCTGCAATTCAGCTCCTCTGTGTTTGGCCGCTGGAT 207
QY 21 TTPAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluLeuGlyTyrSerTyr 40
Db 208 TGGCTTAATGGATCTACAGACACAGAGAAACTTGTGGAAGATTGGCTGGTCTCTAT 267
QY 41 ThrGlyAlaLeuAsnGlnLysAsnTyrGlyLysLysTyrProThrCysAsnSerProLys 60
Db 268 ACAGGAGCACTGAATCAAAAAATTTGGGAAAGAAATATCCAACATGTATAGCCCAAA 327
QY 61 GlnSerProIleAsnLeuLeuLeuLeuLeuLeuLeuValAsnValAsnLeuLysLysLeu 80
Db 328 CAATCTCTCTATCAATATGATGAAGATCTTACAGATAAATGTGAATCTTAAGAAACTT 387
QY 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db 388 AAATTTCCGGTGGGATAAAACATCATTTGGGAAAACACATTCATTCAATCAACTGGGAA 447
QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyValSerGluMetVal 120
Db 448 ACAGTGGAAATTAATCTCACTAATGACTACCGTGTGAGGAGGAGTTTCAGAAATGGTG 507
QY 121 PheLysAlaSerLysIleThrPheHisTyrGlyLysCysAsnMetSerSerAspGlySer 140
Db 508 TTTAAGCAAGCAAGATAACTTTTCTCGGGAAATGCAATATGTCATCTCATGGATCA 567
QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 568 GAGCATAGTTTAGAAGGACAAAAATTTCCATTTGAGATGCAATCTACTGCTTTGATGCA 627
QY 161 AspArgPheSerSerPheGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 628 GACCCGATTTCAAGTTTTCAGGAGCAGTCAAGAGGAAAGGAGTAAAGAGCTTTATACC 687
QY 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleLeuAspGlyVal 200
Db 688 ATTTGTTTGGAGTTGGGACAGAGAAATTTGGATTTTCAAGCGCATATTATGATGGATG 747
QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
Db 748 GAAAGTTTAGTCTGTTTGGGAGCAGCGTCTTTAGATCCATTCATCTGTTGAACTT 807
QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProCys 240
Db 808 CTGCCAAACTCACTGACAAAGTATTACATTTACAATGGCTCATTTGACATCTCTCCCTGC 867
QY 241 ThrAspThrValAspTyrIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 868 ACGACACAGATTGACTGGATTTGTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 928 GCTGTTTTTGGAGTTCTTACATGCAACATCTGGTTATGTCATGCTGATGACCTAC 987
QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 988 TTACAAAAACAATTTTCGAGAGCAACAGTACAGTTCTCTAGACAGGTGTTTCTCTCATAC 1047
QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 1048 ACTGGAAGGAGAGATTCATGAAGCAGTTGTAGTTTCAAGAACCAAGAAATGTTCCAGCT 1107
QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTyrGluArgProArgValValTyrAsp 340
Db 1108 GACCCAGAGATTATACAGCTCTTCTTACATGCGGAAGACCTCGAGTCTGTTATGAT 1167
QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 1168 ACCATGATTGAGAGTTTGCAGTTTGTACAGCAGTTTGGATGGAGAGACCAACCAAG 1227
QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db 1228 CATGAATTTTTCACAGATGGCTATCAAGACTTGGGTGCTATT 1269

RESULT 8

ADN39057
ID ADN39057 standard; cDNA; 5343 BP.
XX
AC ADN39057;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:375.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
(E0SB-) EOS BIOTECHNOLOGY INC.
PA
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR WPI; 2003-468649/44.
DR P-FSDB; ADN39058.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 8; SEQ ID NO 375; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC nucleic acid also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,

CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularization syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.

XX
 SQ Sequence 5343 BP; 1706 A; 1058 C; 1092 G; 1487 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 7,17e-206 Length: 5343
 Score: 1968.00 Matches: 372
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 2
 Query Match: 99.39% Indels: 0
 DB: 11 Gaps: 0

US-09-983-000a-2 (1-374) x ADN39057 (1-5343)

Qy 1 MetArgIleLeuLysAtrPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
 Db 148 ATGGAAATCCTTAAGCGTTTCTGCTTGCAATTCAGCTCCTCTGTTGTCGGCTGGAT 207
 Qy 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40
 Db 208 TGGGCTAATGGATACACAGACACAGAGAAACCTTGTTGAAGAGATTGGCTGGCTCTAT 267
 Qy 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
 Db 268 ACAGGAGCACTGAATCAAAAATTTGGGGAAGAAATATCAACATGTAATAGGCCCAAAA 327
 Qy 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
 Db 328 CAATCTCTATCAATATTGATGAGATCTTACACAGTAATGTAATCTTAAGAACTT 387
 Qy 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
 Db 388 AAAATTCAGGGTTGGGATAAACATCATTTGGAAAAACACATTCATTAACACCTGGGAAA 447
 Qy 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyValSerGluMetVal 120
 Db 448 ACAGTGGAAATTAATCTCACAATGACTACCGTGTGACGGAGGATTTTCAGAAATGGTG 507
 Qy 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
 Db 508 TTTAAAGCAGCAGATTAACCTTTTCACTGGGAAAAATGCAATATGTCATCTGATGATCA 567
 Qy 141 GluIleSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
 Db 568 GAGATAGTTTAGAAGGACAAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATGCA 627
 Qy 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
 Db 628 GACCGATTTTCAAGTTTTTGGAGAGCAGTCAAGGAAAGGAAAGTTTAAGAGCTTTATCC 687
 Qy 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIleAspGlyVal 200
 Db 688 ATTTGTTTGGATTGGGACAGAGAAATTTGGATTTCAAGCGATTTATTGATGGATC 747
 Qy 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspPropheIleLeuLeuAsnLeu 220
 Db 748 GAAAGTGTAGTCTGTTTGGGAAGCAGGCTGCTTTAGATCCATTCAATCTGTTGAACCTT 807
 Qy 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240
 Db 808 CTGCCAAACTCAACTGCAAGTATTACATTTTACATTTGCTTATGATGCTCTCTCCCTGC 867
 Qy 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
 Db 868 ACAGACACAGTTGACTGATGTTTATTAAGATACAGTTAGCACTCTGGAAGCCAGTTG 927
 Qy 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
 Db 928 GCTGTTTTTTGTGAAGTCTTACATGCAACAATCTGTTTATGTCTATGCTGATGACTAC 987

Qy 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
 Db 988 TTACAAACAATTTTCAGAGACACAGTACAGTTCTCTAGACAGGTGTTTCTCTCATAC 1047
 Qy 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
 Db 1048 ACTGGAAGGAAGAGATTTCATGAAGCAGTTTGTAGTTTCAGAACCCAGAAATGTTCCAGCT 1107
 Qy 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
 Db 1108 GACCCAGAGAAATTATACACAGCTTCTTTGTATACATGGGAAAGACCTCGAGTCGTTTATGAT 1167
 Qy 341 ThrMetIleGluLysPheAlaValIleTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
 Db 1168 ACCATGATTGAGAAGTTTGACAGTTTGTACACAGATTTGTACAGAGTTGGATGGAGGACCAACCAAG 1227
 Qy 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
 Db 1228 CATGAATTTTGACAGATGGCTATCAAGACTTGGGTGCTATT 1269
 RESULT 9
 ID ABX76417
 XX ABX76417 standard; DNA; 5367 BP.
 AC ABX76417;
 XX 02-APR-2003 (first entry)
 DT Lung cancer-associated polynucleotide #281.
 DE Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX Unidentified.
 OS
 XX WO200286443-A2.
 FN 31-OCT-2002.
 PD 18-APR-2002; 2002WO-US012476.
 PF 18-APR-2001; 2001US-0284770P.
 PR 10-MAY-2001; 2001US-0290492P.
 PR 09-NOV-2001; 2001US-0339245P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0334370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PA Aziz N, Murray R;
 PI WPI; 2003-093161/08.
 DR P-PSDB; ABUS6688.
 DR
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX
 FS Claim 22; Page 404-405; 453pp; English.
 XX
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridises
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung

CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention

XX
 SQ Sequence 5367 BP; 1714 A; 1063 C; 1098 G; 1492 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7,21e-206 Length: 5367
 Score: 1968.00 Matches: 372
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 2
 Query Match: 99.39% Indels: 0
 DB: 8 Gaps: 0

US-09-983-000A-2 (1-374) x ABX76417 (1-5367)

QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
 DB 148 ATGCGAATCCTAAAGCGTTTCTCGCTTCGATTAGCTTCTGCTGTTTGGCGCTGGAT 207
 QY 21 TpdAlaAenGlyTyrArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
 DB 208 TGGGCTTAATGGGATCTACAGACACAGAGAAACTTGTTCAGAGATTGGCTGGTCTTAT 267
 QY 41 ThrGlyValLeuAenGlnLysAenTrpGlyLysLysTyrProThrCysAenSerProLys 60
 DB 268 ACAGAGACACTGAATCAAAAATAATTTGGGAGAGAAATATCCACATGTAATAGCCCAAAA 327
 QY 61 GlnSerProIleAenLysAspLeuThrGlnValAenValAenLysLysLeu 80
 DB 328 CAATCTCTATCAATATGATGAGATCTTACACAGTAATATGTAATGTAATGTAATGTAAT 387
 QY 81 LysPheGlnGlyTrpAspLysThrSerLeuGluAenThrPheIleAenThrGlyLys 100
 DB 388 AATTTTCAGGGTGGGTAATAACATCATTTGGGAAACACATCATTCATTAACACTGGGAAA 447
 QY 101 ThrValGluIleAenLysAenAspTyrArgValSerGlyValSerGluMetVal 120
 DB 448 ACAGTGGAAATTAATCTCAATATGATGAGTACCTGTACAGGGAGGATTTTCAAGATGGTG 507
 QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAenMetSerSerAspGlySer 140
 DB 508 TTTAAAGCAAGCAAGATAAATTTTCACTGGGAAATGCAATATGTCATCTGATGATCA 567
 QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
 DB 568 GAGCATAGTTTAGAAGGACAAAATTTCCACTGTGAGATGCAATCTACTGCTTTGATGCG 627
 QY 161 AspArgPheSerSerPheGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
 DB 628 GACCGATTTTCAAGTTTGGAGAGACAGTCAAGAGAAAGGAGTTTAAAGCTTTATTC 687
 QY 181 IleLeuPheGluValGlyThrGluGluAenLeuAspPheLysAlaIleIleAspGlyVal 200
 DB 688 ATTTTGTGTAGTTTGGGACAGAGAAATTTGGATTTTCAAGCGATTTATTGATGGAGTC 747
 QY 201 GluSerValSerArgPheGlyGlnAlaLeuAenLeuAspPheIleLeuLeuAenLeu 220
 DB 748 GAAAGTGTGTAGTGGTGGGAGCAGGCTGCTTTAGATCCATTCATTCATTCATTCATTCAT 807
 QY 221 LeuProAenSerThrAspLysTyrTrpIleTyrAenGlySerLeuThrSerProProCys 240
 DB 808 CTGCCAACTCACTGACAGATTTACATTTTACATGCTCTTATGACATCTCTCCCTGTC 867
 QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260

DB 868 ACAGACACAGTTGACTGATTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
 QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
 DB 928 GCTGTTTTTGTGAAGTTCTTTACCAATGCAACATCTGTTATGTCATGCTGATGACTAC 987
 QY 281 LeuGlnAenAenPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
 DB 988 TTACAAACAATTTTCGAGAGACACAGTCAAGTTCTCTAGACAGGTGTTTCTTCATAC 1047
 QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAenValGlnAla 320
 DB 1048 ACTGGAAGGAGAGATTCATGAAGCAGTTTGTAGTTTCTAGAACACAGAAATGTTCCAGCT 1107
 QY 321 AspProGluAenTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
 DB 1108 GACCCAGAGAAATATATACAGCTTCTTTGTTATCATGGGAAAGACCTCGAGTCTGTTATGAT 1167
 QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
 DB 1168 ACCATGATTGAGAGTTTGCAGTTTGTACACGAGTTTGGATGGAGAGACCAAAACCAAG 1227
 QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
 DB 1228 CATGAATTTTTCAGACAGATGGCTATCAAGACTTGGTGCTATT 1269
 RESULT 10
 ID ABX76218 standard; DNA; 5367 BP.
 XX AC ABX76218;
 XX DT 02-APR-2003 (first entry)
 XX DE Lung cancer-associated polynucleotide #87.
 XX KW Lung cancer-associated polynucleotide; gene; da; cytostatic; emphysema;
 KW anti-inflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX OS Unidentified.
 XX PN WO200286443-A2.
 XX FD 31-OCT-2002.
 XX PF 18-APR-2002; 2002WO-US012476.
 XX PR 18-APR-2001; 2001US-0284770P.
 PR 10-MAY-2001; 2001US-0290492P.
 PR 09-NOV-2001; 2001US-0339245P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0334370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX AZ Aziz N, Murray R;
 XX WI; 2003-093161/08.
 DR P-FSD; ABUS6494.
 XX
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX Claim 22; Page 258-259; 453pp; English.
 XX The invention relates to a method for detecting a lung cancer-associated

CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridises
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention

XX
 SQ Sequence 5367 BP; 1714 A; 1063 C; 1098 G; 1492 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7,21e-206 Length: 5367
 Score: 1968.00 Matches: 372
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 2
 Query Match: 99.39% Indels: 0
 DB: 8 Gaps: 0

US-09-983-000A-2 (1-374) x ABX76218 (1-5367)

QY 1 MetArgIleLeuLysArgPheIleuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
 DB 148 ATGCAATCCATAAGCGTTTCTCTGCTTGCATTACGCTCTCTGTGTTCGCGCTGGAT 207
 QY 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40
 DB 208 TGGGCTAATGATGACTACAGACACACAGAGAAACTTGTGTGAAGAGATGGCTGTCTTAT 267
 QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
 DB 268 ACAGGAGCACTGAATCAAAAATTTGGGAAAGAAATATCAACATGTAATAGCCCAAAA 327
 QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLeu 80
 DB 328 CAATCTCTATCAATATTGATGAAGATCTTACACAGTAATATGTAATCTTAAAGAACTT 387
 QY 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
 DB 388 AAATTTCAAGGTTGGGATTAACAATCATTTGGGAAACACATTCATTCATACACTGGGAAA 447
 QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyValSerGluMetVal 120
 DB 448 ACAGTGGAAATTAATCTCACTAATGACTACCGTGTGACGGGAGGATTTCAAGAAATGGTG 507
 QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
 DB 508 TTTAAAGCAAGCAAGATAACTTTTTCACCTGGGGAAATGCAATATGTCATCTGATGGATCA 567
 QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
 DB 568 GAGCATAGTTTAGAAGGACAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATGGC 627
 QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
 DB 628 GACCGATTTTCAAGTTTTCAGAGAGACAGTCAAGAGAAAGGAAAGTAAAGAGCTTTATCC 687
 QY 181 IleLeuPheGluValGlyThrGluLeuAsnLeuAspPheLysAlaIleIleAspGlyVal 200
 DB 688 ATTTTGTGTGAGTTGGGACAGAGAAATTTGGATTTCAAGCGCATTTATTGATGGAGTC 747
 QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220

DB 748 GAAAGTCTTAGTCTGTTTTGGGAAGCAGCGTCTTTAGATCCATTCATCATCTGTTGAACCTT 807
 QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240
 DB 808 CTGCCAAACTCAACTGACAAAGTATTACATTTACAATGGCTCATTTGACATCTCTCTCCCTGC 867
 QY 241 ThrAspThrValAspTyrIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
 DB 868 ACAGACACAGTTGACTGGATTTGTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
 QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
 DB 928 GCTGTTTTTGTGAGTCTTACAAATGCAACATCTGTTATGTCATGCTGATGACTAC 987
 QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
 DB 988 TTACAAACAATTTTCGAGAGCAACAGTCAAGTTCCTAGACAGGTGTTTTTCTCTCATAC 1047
 QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
 DB 1048 ACTGGAAGGAAGAGATTTCATGAAGCAGTTGTAGTTTCAGAACCAAGAAATGTCAGCT 1107
 QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
 DB 1108 GACCCACAGAAATATACACAGCTTCTTGTGTACATGGGAAAGACCTCGAGTCGTTTATGAT 1167
 QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
 DB 1168 ACCATGATTGAGAAGTTTGCAAGTTTTGTACCAGCAGTTGGATGGAGAGAGCAACCAAG 1227
 QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
 DB 1228 CATGAATTTTTCAGCATGGCTATCAAGACTTGGGTGCTATT 1269
 RESULT 11
 ID ADN39055 standard; cDNA; 5367 BP.
 XX ADN39055;
 AC ADN39055;
 DT 17-JUN-2004 (first entry)
 XX
 DB Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:373.
 XX Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.
 XX Homo sapiens.
 OS
 PN WO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 PF 13-NOV-2002; 2002WO-US036810.
 XX
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.

PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 PA (E08B-) EOS BIOTECHNOLOGY INC.
 PI Afar D, Aziz N, Gineburg WM, Gish KC, Glynn R, Revezzi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KB, Zlotnik A;
 XX
 DR WPI: 2003-468649/44.
 DR P-PSDB; ADN39056.
 XX
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 8; SEQ ID NO 373; 1385pp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX
 SQ Sequence 5367 BP; 1714 A; 1063 C; 1098 G; 1492 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 7,21e-206 Length: 5367
 Score: 1968.00 Matches: 372
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 2
 Query Match: 99.39% Indels: 0
 DB: 11 Gaps: 0
 US-09-983-000A-2 (1-374) x ADN39055 (1-5367)
 QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuLysValCysArgLeuAsp 20
 DB 148 ATGCGAATCTTAAGCGTTTCTCGCTTGCAATCAGCTCTCTGTGTGGCGCTGGAT 207
 QY 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
 DB 208 TGGGCTAATGGATACACAGACCAAGAGAAATCTGTGTGAAGAGATTGGCTGCTTAT 267
 QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
 DB 268 ACAGGAGCACTGAATCAAAAATAATTTGGGGAAGAAATATCCAAACATGTAATAGCCCAAA 327
 QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
 DB 328 CAATCTCTATCATATTGATGAGATCTTACACAAAGTAATGTGAATCTTAAGAAACCTT 387
 QY 81 LysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
 DB 388 AAATTTTCAGGTTGGGATAAACAATCATTTGGAACAACATCATTCATTAACACTGGGAAA 447
 QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyLysValSerGluMetVal 120
 DB 448 ACAGTGGAAATTAATCTCACTAATGACTACCGTGTGAGCGGGAGGAGTTTCAGAAATGGTG 507

QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
 DB 508 TTTAAAGCAGCAAGATAAATCTTTTCACTGGGGAATAATGATATGTCACTCTGATGGATCA 567
 QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
 DB 568 GAGCATAGTTTAGAAGGAGCAAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATGCG 627
 QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
 DB 628 GACCGATTTTCAAGTTTGGAGAGCAGTCAAGAGAAAGGAGTAAAGAGCTTTATGCC 687
 QY 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleAspGlyVal 200
 DB 688 ATTTGTTTGGAGTTGGGACAGAGAAAATTTGGATTTCAAAGCGATTTATTGATGGAGTC 747
 QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspPropheIleLeuLeuAsnLeu 220
 DB 748 GAAAGTGTTAGTCGTTTGGGAGCAGCTGCTTTAGATCCATTCATCTGTTGAACCTT 807
 QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrHisGlySerLeuThrSerProProCys 240
 DB 808 CTGCCAAACTCACTGACAGATATTACATTTACAAATGGCTCATTTGACATCTCCTCCCTGC 867
 QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
 DB 868 ACAGACACAGTTGACTGGATGTTTTTAAAGATACAGTAGTCATCTCTGAAAGCCAGTTG 927
 QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
 DB 928 GCTGTTTTTGTGAAGTCTTTACAAATGCAACAATCTGCTTATGTCATGCTGATGACTAC 987
 QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
 DB 988 TTACAAACAATTTTCGAGAGCAACAGTACAAAGTCTCTAGACAGGTGTTTTCTCATAC 1047
 QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
 DB 1048 ACTGGAAGGAGAGAGATTCATGAAGCAGTTGTAGTTCAGAACCCAGAAATGTTCAAGCT 1107
 QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
 DB 1108 GACCCAGAGAAATTAACAGCTTCTTGTGTACATGGGGAAGACCTCGAGTCTGTTTATGAT 1167
 QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
 DB 1168 ACCATGATTGAGAAAGTTTGCAGTCTTTGTACCAGCAGTTGGATGGAGAGACCAACCAAG 1227
 QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
 DB 1228 CATGAAATTTTGCACAGATGGCTATCAAGACTTTGGGTGCTATT 1269
 RESULT 12
 ADN39726
 ID ADN39726 standard; cDNA; 5367 BP.
 XX
 AC ADN39726;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:C98.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.
 XX
 OS Homo sapiens.
 XX

PN W02003042661-A2.
 XX PD 22-MAY-2003.
 XX PF 13-NOV-2002; 2002WO-US036810.
 XX PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 09-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX WPI: 2003-468649/44.
 DR P-PSDB; ADN39943.
 XX
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 8; SEQ ID NO C98; 1385pp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the;
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX
 SQ Sequence 5367 BP; 1714 A; 1063 C; 1098 G; 1492 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 7, 21e-206 Length: 5367
 Score: 1968.00 Matches: 372
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 2
 Query Match: 99.39% Indels: 0
 DB: 11 Gaps: 0
 US-09-983-000A-2 (1-374) x ADN39726 (1-5367)
 QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
 DB 148 ATGCGAATCCTAAAGCGTTCTCGCTGCATTCAGCTCCTCTGTTGCGCGCTGGAT 207

QY 21 TTPAlaAenGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTTPSerTyr 40
 DB 208 TGGGCTAATGGATACACAGAACTTGTGTTGAAGAGATTGGTGTCTTAT 267
 QY 41 ThrGlyAlaLeuAenGlnLysAenTTPGlyLysLysTyrProThrCysAenSerProLys 60
 DB 268 ACAGGAGCACTGAATCAAAAAAATTTGGGAAAGAAATATCAACATGTAATAGCCAAAA 327
 QY 61 GlnSerProIleAenIleAspGluAAspLeuThrGlnValAenValAenLeuLysLysLeu 80
 DB 328 CAATCTCTATCAATATTGATGAAGATCTTACACAAGTAATGTGAATCTTAAGAAAACTT 387
 QY 81 LysPheGlnGlyTTPAspLysThrSerLeuGluAenThrPheIleHisAenThrGlyLys 100
 DB 388 AAATTTGAGGGTGGGATAAAACATCATTTGGAANAACATTCATTTAATACACTGGGAAA 447
 QY 101 ThrValGluIleAenLeuThrAenAspTyrArgValSerGlyGlyValSerGluMetVal 120
 DB 448 ACAGTGGAAATTAATCTCACTAATGACTACCGTGTACGGGAGGAGTTTCAGAAATGGTG 507
 QY 121 PheLysAlaSerLysIleThrPheHisTTPGlyLysCysAenMetSerSerAspGlySer 140
 DB 508 TTTAAAGCAAGCAAGATAACTTTTCACTGGGAAAAATGCAATATGTCTCATCTGATGGATCA 567
 QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
 DB 568 GAGCATAGTTTAGAAGGACAAAAATTTCCACTTGAGATGCNAATCTACTGCTTTGATGCG 627
 QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
 DB 628 GACCGATTTTCAAGTTTGGAGGAGCAGTCAAAGGAAAAAGGAAAGTTTAAGAGCTTTATCC 687
 QY 181 IleLeuPheGluValGlyThrGluGluAenLeuAspPheLysAlaIleIleAspGlyVal 200
 DB 688 ATTTTGTGTTGAGGTGGGACAGAAATAATTTGGATTTCAAAGCGATTATTATGATGGATC 747
 QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
 DB 748 GAAAGTGTGTAGTCTTTTGGAGAGCAGCTGCTTTAGATCAATTCATCTGTTGAACTT 807
 QY 221 LeuProAenSerThrAspLysTyrTyrIleTyrAenGlySerLeuThrSerProProCys 240
 DB 808 CTGCCAAACTCAACTGACAGATATTACATTTACAAATGGCTCAATTGACATCTCTCCCTGTC 867
 QY 241 ThrAspThrValAspTTPIleValPheLysAspThrValSerIleSerGluSerGluLeu 260
 DB 868 ACAGACACAGTTGACTGCTGATTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
 QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
 DB 928 GCTGTTTTTTGTGAAGTTCTTACAAATGCACAACTCTGTTATGTCATGCTGATGACTAC 987
 QY 281 LeuGlnAenAenPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
 DB 988 TTCAAAACAATTTTCGAGAGCAACAGTACAAGTTCTCTAGACAGGTGTTTTCTCTCATAC 1047
 QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAenValGlnAla 320
 DB 1048 ACTGGAAGGAAGAGATTCATGAAGCAGTTGTGTAGTTTCAGAACCCAGAAAAATGTTGAGCT 1107
 QY 321 AspProGluAenTyrThrSerLeuLeuValThrTTPGluArgProArgValValTyrAsp 340
 DB 1108 GACCCAGAGAAATTAATACAGCTTCTTGTGTACATGGGAAGACCTCGAGTCTGTTATGAT 1167
 QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
 DB 1168 ACCATGATTGAGAAAGTTTGCAGTTTGTACCAGCAGTTGGATGGAGAGGACCAACCAAG 1227
 QY 361 HisGluPheLeuThrAspGlyTyrGlnAAspLeuValThrIle 374
 DB 1228 CATGAATTTTTCAGACAGATGGCTATCAAGACTTTGGGTGCTATT 1269
 RESULT 13

ABX76420

ID ABX76420 standard; DNA; 5481 BP.

AC ABX76420;

XX 02-APR-2003 (first entry)

XX Lung cancer-associated polynucleotide #284.

KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

OS

XX WO200286443-A2.

XX

PD 31-OCT-2002.

PF 18-APR-2002; 2002WO-US012476.

XX

PR 18-APR-2001; 2001US-0284770P.

PR 10-MAY-2001; 2001US-0290492P.

PR 09-NOV-2001; 2001US-0339245P.

PR 13-NOV-2001; 2001US-0350666P.

PR 29-NOV-2001; 2001US-0334370P.

PR 12-APR-2002; 2002US-0372246P.

XX

PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX

PI Aziz N, Murray R;

XX

DR WPI; 2003-093161/08.

DR P-PSDB; ABUS6691.

XX

PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.

XX

PS Claim 22; Page 409-410; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention

XX Sequence 5481 BP; 1734 A; 1104 C; 1123 G; 1520 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7, 44e-205 Length: 5481
 Score: 1968.00 Matches: 372
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 2
 Query Match: 99.39% Indels: 0
 DB: 8 Gaps: 0

US-09-983-000A-2 (1-374) x ABX76420 (1-5481)

QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
 DB 148 ATGCGAATCTCTAAACCGTTTCCTCGCTTGCATTGAGTCTCTCTGTGTGTGGCGCTGGAT 207

QY 21 TrpAlaLeuGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40
 DB 208 TGGGCTAATGGTACTACTACAGACAACAGAGAAACTTGTGTGAAGAGATTGGCTGGTCTTAT 267

QY 41 ThrGlyAlaLeuLeuGlnLysAsnTrpGlyLysTyrProThrCysAsnSerProLys 60
 DB 268 ACAGGAGCACTGAATCAAAAAAATGGGAAAGAAATATCAACATGTAAATAGGCCAAAA 327

QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLeu 80
 DB 328 CAATCTCTCTATCAATATTTGATGAAGATCTTTACACAGTAAATGTGAATCTTTAAGAAACTT 387

QY 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
 DB 388 AATTTTCAAGGTTGGGATATAAACATCATTTGGAAACACATTCATTCATTAACACTGGGAAA 447

QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
 DB 448 ACAGTGGAAATTAATCTCACTAATGACTACCGTGTACGCGGAGGAGTTTCAGAAATGGTG 507

QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
 DB 508 TTTAAAGCAAGCAAGATAACTTTTCTCTGGGAAATATGCAATATGTCTCATCTGTATGATCA 567

QY 141 GluHisSerLeuGluGlyGlnLysPhePheProLeuGluMetGlnIleTyrCysPheAspAla 160
 DB 568 GAGCATAGTTTAGAAGGACAAAAATTTCCACTTGAGATGCAAACTCTACTGCTTTGATGCG 627

QY 161 AspArgPheSerSerPheGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
 DB 628 GACCGATTTTCAAGTTTTGAGGAGCAGTCAAGGAAAGGAAAGTAAAGAGCTTTATACC 687

QY 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIleAspGlyVal 200
 DB 688 ATTTTGTGTGAGGTTGGGACAGAGAAATTTGGATTTCAAGCGATTTATTGTATGGAGTC 747

QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspPropheIleLeuLeuAsnLeu 220
 DB 748 GAAAGTGTAGTCTGTTTGGGAGCAGCTGCTTTAGATCCTATTCATCTGTTGAACTT 807

QY 221 LeuProHisSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240
 DB 808 CTGCCAAACTCAACTGACCAAGTATTACATTTACAATGGCTCATTTGACATCTCTCCCTGCG 867

QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
 DB 868 ACAGACACAGTTGACTGGATTTGTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927

QY 261 AlaValPheCysGluValLeuLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
 DB 928 GCTGTTTTTGTGAAGTCTTTACAAATGCAAACTCTGCTTATGTCATGCTCATGAGCTAC 987

QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
 DB 988 TTACAAAACAAATTTTCGAGAGCAACAGTACAGATTCTCTAGACAGGTGTTTTCTCATAC 1047

QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
 DB 1048 ACTGGAAGGAAGAGAGATTTCATGAAGCAGTTTGTAGTTTCAGAAACAGAAAAATGTTTCAGGCT 1107

QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
 DB 1108 GACCCAGAGAAATATACAGCTTCTTGTGTACATGGGAAAGACCTCTGAGTGTGTTATGAT 1167

QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
 DB

Db 1168 ACCATGATTGAGAAAGTTTGCAGTTTTTTGTACAGCAGTTGGATGGAGAGACCAAAACCAAG 1227
 Qy 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
 Db 1228 CATGAATTTTTCAGAGATGGCTATCAAGACTTGGTGCTATT 1269
 RESULT 14
 ABX76221
 ID ABX76221 standard; DNA; 5481 BP.
 XX
 AC ABX76221;
 XX
 DT 02-APR-2003 (first entry)
 XX
 DE Lung cancer-associated polynucleotide #90.
 XX
 KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 OS Unidentified.
 OS
 XX
 PN W0200286443-A2.
 XX
 PD 31-OCT-2002.
 XX
 PF 18-APR-2002; 2002WO-US012476.
 XX
 PR 18-APR-2001; 2001US-0284770P.
 PR 10-MAY-2001; 2001US-0290492P.
 PR 09-NOV-2001; 2001US-0339245P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0334370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Aziz N, Murfay R;
 XX
 WPI; 2003-091161/08.
 DR P-PSDB; ABUS4497.
 DR
 XX
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX
 PS Claim 22; Page 262-263; 453pp; English.
 XX
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridises
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX
 SQ Sequence 5481 BP; 1734 A; 1104 C; 1123 G; 1520 T; 0 U; 0 Other;

QY 321 AppProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
 DB 1108 GACCAGAGAAATTATACAGCCCTCTGTGTACATGGGAAGACCTCGAGTCGTTATGAT 1167
 QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
 DB 1168 ACCATGATTTGAGAGCTTTGTCAGTTTGTACACGACGTTGGATGGAGAGCAACCAACG 1227
 QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
 DB 1228 CATGATTTTACAGATGGCTATCAAGACTTGGGTGCTATT 1269
 RESULT 15
 ADN39061
 ID ADN39061 standard; cDNA; 5481 BP.
 XX
 AC ADN39061;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:379.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnerability; gene therapy; vaccine; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 PF 13-NOV-2002; 2002WO-US036810.
 XX
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 04-APR-2002; 2002US-0368809P.
 PR 12-APR-2002; 2002US-0370110P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-039775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 DR WPI; 2003-468649/44.
 DR P-FSDB; ADN39062.
 XX
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 8; SEQ ID NO 379; 1385pp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)

CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.

XX
 SQ Sequence 5481 BP; 1734 A; 1104 C; 1123 G; 1520 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7,44e-206 Length: 5481
 Score: 1968.00 Matches: 372
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 2
 Query Match: 99.39% Indels: 0
 DB: 11 Gaps: 0

US-09-983-000A-2 (1-374) x ADN39061 (1-5481)

QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
 DB 148 ATGCGAATCCTAAACCGTTTCTCGCTTGCATTGAGCTCTCTGTGTGGCGCTGGAT 207
 QY 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
 DB 208 TGGCTAATGGATACTACAGACACAGAGAAAACCTTGTGGAAGAGATGGCTGTGCTTAT 267
 QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
 DB 268 ACAGGAGCACTGAATCAAAAAAATTTGGGAAAGAAATATCCACATGTATATAGCCCAAAA 327
 QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
 DB 328 CAATCTCCTATCAATATTGATGAAGATCTTACACAGTAAATGTGAATCTTAAAGAACTT 387
 QY 81 LysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
 DB 388 AAATTTCCAGGGTTGGGATAAACAATCATTTGGAAACACATTCATTATACACACCTGGGAAA 447
 QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
 DB 448 ACAGTGGAAATTAATCTCACTAATGACTACCGTGTACGGGAGGAGTTTCAGAAATGGTG 507
 QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
 DB 508 TTTAAAGCAAGCAAGATAAATCTTTTCACTGGGAAAAATGCAATATGTCTATCTGATGTGATCA 567
 QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
 DB 568 GAGCATAGTTTAGAAGGACAAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTTGATGCG 627
 QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
 DB 628 GACCGATTTTCAGTTTTCAGTTTTCAGTTTTCAGTTTTCAGTTTTCAGTTTTCAGTTTTCAG 687
 QY 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIleAspGlyVal 200
 DB 688 ATTTGTTGTTGAGGTTGGGACAGAAAGAAATTTGCAATTTCAAGCGATTAATTGATGGAGTC 747
 QY 201 GluSerValSerArgPheGlyLysGlnAlaAlaLeuAspPhePheIleLeuLeuAsnLeu 220
 DB 748 GAAAGTGTGTAGTCGTTTGGGAAAGCAGGCTGCTTTAGATCCATTCATCTACTGTTGAACCTT 807

```
Qy 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240
Db |||||
Db 808 CTGCCAAACTCAACTGACAAAGTATTACATTTACAATGGCTCATTGACATCTCCTCCCTGC 867
Qy 241 ThrAspThrValAspTyrIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db |||||
Db 868 ACAGACACAGTTGACTGGATTGTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
Qy 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db |||||
Db 928 GCTGTTTTTGTGAAGTTCTTACAATGCACAATCTGGTTATGTCAATGCTGATGGACTAC 987
Qy 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db |||||
Db 988 TTACAAAAACAATTTTCGAGAGCAACAGTACAAAGTTCTCTAGACAGGTGTTTTTCCTCATAC 1047
Qy 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db |||||
Db 1048 ACTGGAAAGGAGAGATTCTAGNAGCAGTTTGTAGTTCAGAACCCAGAAAATGTTCAGGCT 1107
Qy 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
Db |||||
Db 1108 GACCCAGAGAATTATACCAGCCTTCTTGTTACATGGGAAAGACCTCGAGTCGTTTATGAT 1167
Qy 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrIys 360
Db |||||
Db 1168 ACCATGATGTGAGAAAGTTTGCAGTTTGTACCAGCAGTTGGATGGAGAGGCCAAACCAAG 1227
Qy 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db |||||
Db 1228 CATGAATTTTTCACAGATGGCTATCAAGACTTGGGTGCTATT 1269
```

Search completed: August 30, 2005, 09:58:15
Job time : 593 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 30, 2005, 09:48:33 ; Search time 689 Seconds
(without alignments)
3551.692 Million cell updates/sec

Title: US-09-983-000A-2
Perfect score: 1980
Sequence: 1 MRILRFLACIQLLCVCRLD.....GEDQTKHFLTDGYQDLVTI 374

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgm2_1/USPTO_spool_p/US09983000/runat_29082005_155926_14737/app_query.fasta_1.519
-DB=PublishedApplications NA -Qfmt=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum2
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09983000 @CGN 1 1 723 @runat_29082005_155926_14737
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgm2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgm2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgm2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgm2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgm2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgm2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgm2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgm2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgm2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgm2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgm2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgm2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgm2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgm2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgm2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgm2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgm2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgm2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgm2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgm2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgm2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
23: /cgm2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
24: /cgm2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgm2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgm2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1980	100.0	3091	10	US-09-983-000A-1	Sequence 1, Appli
2	1980	100.0	3091	21	US-10-652-981-1	Sequence 1, Appli
3	1968	99.4	3340	17	US-10-295-027-1047	Sequence 1047, Ap
4	1968	99.4	4220	17	US-10-295-027-1049	Sequence 1049, Ap
5	1968	99.4	4347	17	US-10-295-027-1042	Sequence 1042, Ap
6	1968	99.4	4480	17	US-10-295-027-1048	Sequence 1048, Ap
7	1968	99.4	5343	17	US-10-295-027-375	Sequence 375, App
8	1968	99.4	5367	17	US-10-295-027-373	Sequence 373, App
9	1968	99.4	5367	17	US-10-295-027-1044	Sequence 1044, Ap
10	1968	99.4	5481	17	US-10-295-027-379	Sequence 379, App
11	1968	99.4	7941	9	US-09-954-456-1812	Sequence 1812, Ap
12	1968	99.4	7941	9	US-09-816-703A-1	Sequence 1, Appli
13	1968	99.4	7941	10	US-09-983-000A-5	Sequence 5, Appli
14	1968	99.4	7941	17	US-10-295-027-381	Sequence 381, App
15	1968	99.4	7941	21	US-10-843-641A-4839	Sequence 4839, Ap
16	1968	99.4	7941	21	US-10-652-981-5	Sequence 5, Appli
17	1968	99.4	7941	21	US-10-936-626-32	Sequence 32, Appl
18	1968	99.4	7941	21	US-10-938-061-32	Sequence 32, Appl
19	1968	99.4	7944	17	US-10-295-027-371	Sequence 371, App
20	1968	99.4	7944	17	US-10-295-027-1043	Sequence 1043, Ap
21	1968	99.4	7944	21	US-10-852-335A-6	Sequence 6, Appli
22	1968	99.4	8058	10	US-09-983-000A-3	Sequence 3, Appli
23	1968	99.4	8058	21	US-10-652-981-3	Sequence 3, Appli
24	1968	99.4	8292	20	US-10-723-860-5522	Sequence 5522, Ap
25	1959	98.9	3350	20	US-10-723-860-950	Sequence 950, App
26	1955	98.9	5481	17	US-10-295-027-1046	Sequence 1046, Ap
27	1940	98.0	5363	17	US-10-295-027-377	Sequence 377, App
28	1940	98.0	5363	17	US-10-295-027-1045	Sequence 1045, Ap
29	1857	93.8	6801	15	US-10-205-219-120	Sequence 120, App
30	1857	93.8	6801	22	US-10-598-190-29	Sequence 29, Appl
31	1831	92.5	6887	22	US-10-598-190-31	Sequence 31, Appl
32	648.5	32.8	5787	21	US-10-887-553A-434	Sequence 434, App
33	648.5	32.8	6687	20	US-10-723-860-5795	Sequence 5795, Ap
34	412	20.8	2771	18	US-10-240-425-381	Sequence 381, App
35	412	20.8	2771	20	US-10-723-860-3224	Sequence 3224, Ap
36	412	20.8	2771	21	US-10-278-698-170	Sequence 170, App
37	412	20.8	2771	21	US-10-278-698-686	Sequence 686, App
38	412	20.8	2771	22	US-10-756-149-3019	Sequence 3019, Ap
39	412	20.8	2775	17	US-10-172-118-575	Sequence 575, App
40	412	20.8	2775	17	US-10-295-027-295	Sequence 295, App
41	412	20.8	2775	18	US-10-342-887-575	Sequence 575, App
42	412	20.8	2775	19	US-10-648-593-87	Sequence 87, Appl
43	412	20.8	3171	10	US-09-814-353-21734	Sequence 21734, A
44	401	20.3	3151	20	US-10-723-860-7224	Sequence 7224, Ap
45	367	18.5	1657	20	US-10-467-903-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-983-000A-1
; Sequence 1, Application US/09983000A
; Publication No. US20030118585A1
; GENERAL INFORMATION:
; APPLICANT: AGY Therapeutics
; APPLICANT: Melcher, Thorsten
; APPLICANT: Mueller, Sabine
; APPLICANT: Chin, Daniel
; TITLE OF INVENTION: USE OF PROTEIN BIOMOLECULAR TARGETS IN THE TREATMENT AND VISUALIZA
; FILE REFERENCE: 263/180 -- Peagleman -- AGY
; CURRENT APPLICATION NUMBER: US/09/983.000A
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 1
; LENGTH: 3091
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(3091)
; OTHER INFORMATION: PTP-zeta SM1 exon 9 variant
; NAME/KEY: VARSPLIC
; LOCATION: (1262)..(1262)
; OTHER INFORMATION: Alternative splice site
; NAME/KEY: misc feature
; LOCATION: (1273)..(3091)
; OTHER INFORMATION: 3' Untranslated Region
; NAME/KEY: CDS
; LOCATION: (148)..(1272)
; OTHER INFORMATION:
US-09-983-000A-1

Alignment Scores:
Pred. No.:      5,42e-219      Length:      3091
Score:          1980.00      Matches:      374
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              10          Gaps:      0

US-09-983-000A-2 (1-374) x US-09-983-000A-1 (1-3091)

QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
DB 148 ATGCGAATCTTAAGCGTTTCTCGTTGCAATTCAGCTCCTCTGTGTTGGCGCTGGAT 207

QY 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40
DB 208 TGGGCTAATGGATACACAGACAAACAGAGAAACTTGTGTAAGAGATTGGCTGGCTCTAT 267

QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProlys 60
DB 268 ACAGGAGCACTGAATCAAAATAAATTTGGGGAAGAAATAATCCAAATGTAATAGCCCAAA 327

QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
DB 328 CAATCTCTATCAATATTGATGAGATCTTACACAGTAATGTAATGTAATCACTTAAGAACTT 387

QY 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
DB 388 AAAATTCAGGGTTGGGATAAAACATCAATTGGAAAAACACATTCATTCATAACACTGGGAAA 447

QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
DB 448 ACAGTGGAAATTAATCTCACATAATGACATCCGTGTACGGGAGGAGTTTCAGAAATGGTG 507

QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
DB 508 TTTAAGCAGCAGATACATTTTCACTGGGGAATAATGCAATATGTCATCTGATGATCA 567

QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
DB 568 GAGCATAGTTTAGAAGGACAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATGCG 627

QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
DB 628 GACCGATTTTCAAGTTTGGAGAGACGTCGAAGGAAAGGAAAGTTTAAGAGCTTTATCC 687

QY 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleAspGlyVal 200
DB 688 ATTTGTTTGGTTGGACAGAGAAAATTTGGATTTCAAGCGATTTATTGATGGAGTC 747

QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
DB 748 GAAAGTGTAGTCGTTTGGGAAGCAGGCTGCTTTAGATCCATTCATCTGTTGAACCTT 807
```

```
QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240
DB 808 CTGCCAAACTCAACTGACAAGTATTACATTTACAAATGGCTCAATGACATCTCTCCCTGC 867

QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
DB 868 ACAGACACAGTTCAGCTGATTTGTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927

QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
DB 928 GCTGTTTTTGTGAAGTTCTTACAATGCAACAATCTGTTTATGTCTGCTGATGACTAC 987

QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
DB 988 TTAACAAACAAATTTTCGAGAGCAACAGTACAGTTCTCTAGACAGGTGTTTCTCTATAC 1047

QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
DB 1048 ACTGGAAGGAGAGAGATTTCATGAAGCAGTTTGTAGTTTCAGAACCCAGAAAATGTTCCAGGCT 1107

QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
DB 1108 GACCCAGAGAAATATATACAGCTTCTTTGTATACATGGGAAAGACCTCGAGTCTGTTATGAT 1167

QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
DB 1168 ACCATGATTGAGAAGTTTGACAGTTTGTACAGAGTTGGATGGAGAGGACCAACCAAG 1227

QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
DB 1228 CATGAATTTTGACAGATGGCTATCAAGACTTGGTAACATATA 1269

RESULT 2
US-10-652-981-1
; Sequence 1, Application US/10652981
; Publication No. US20050074400A1
; GENERAL INFORMATION:
; APPLICANT: AGY Therapeutics
; APPLICANT: Melcher, Thorsten
; APPLICANT: Mueller, Sabine
; APPLICANT: Chin, Daniel
; TITLE OF INVENTION: USE OF PROTEIN BIOMOLECULAR TARGETS IN THE TREATMENT AND VISUALIZ
; FILE REFERENCE: 263/180 -- Peagleman -- AGY
; CURRENT APPLICATION NUMBER: US/10/652,981
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/983,000A
; PRIOR FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3091
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(3091)
; OTHER INFORMATION: PTP-zeta SM1 exon 9 variant
; FEATURE:
; NAME/KEY: VARSPLIC
; LOCATION: (1262)..(1262)
; OTHER INFORMATION: Alternative splice site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1273)..(3091)
; OTHER INFORMATION: 3' Untranslated Region
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (148)..(1272)
; OTHER INFORMATION:
US-10-652-981-1
Alignment Scores:
```

Pred. No.: 5,42e-219 Length: 3091
Score: 1980.00 Matches: 374
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-983-000a-2 (1-374) x US-10-652-981-1 (1-3091)

QY 1 MetArgileLeuLysArgPheLeuAlaCysIleGlnLeuLysCysValCysArgLeuAsp 20
DB ATGCGAATCTTAAGGTTTCTCTGCTGCAATTCAGTCTCTCTGTTTGGCGCTGGAT 207
QY 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
DB TGGGCTTAATGGATACTACAGACACAGAGAAATCTGTTGAAGAGATTGGCTGGTCTTAT 267
QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
DB ACAGGAGCACTGAATCAAAAATTTGGGAAAGAAATATCCAACTGTAATAGCCCAAA 327
QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
DB CAATCTCTATCAATATTGATGAGATCTTACACAGTAATGTAATCTTAAGAACTT 387
QY 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
DB AAATTTTCAGGTTGGGATAAATCATCTTGGAAACACATCTTATTCATAACACTGGAAA 447
QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyValSerGluMetVal 120
DB ACGTGGGAAATTAATCTCAATGACTACCTGTCAGCGGAGGAGTTTCAGAAATGGTG 507
QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
DB TTTAAGCAGCAGATTAATCTTCTGCGGAAATGCAATATGTCATCTGATGGATCA 567
QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
DB GAGCATAGTTTGAAGGACAAAATTTCCACTTGAGATGCAATCTACTGCTTTGATGGG 627
QY 161 AspArgPheSerSerPheGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
DB GACCGATTTTCAAGTTTGGAGAGCAGTCAAGGAAAGGAGGAGTTAAGAGCTTTATCC 687
QY 181 IleLeuPheGluValGlyThrGluAsnLeuAspPheLysAlaIleLeuAspGlyVal 200
DB ATTTGTTTGGTTGGGACAGAGAAATTTGGATTTTCAAGCGCATTTATGATGGAGTC 747
QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
DB GAAAGTGTAGTCTGTTTGGGAGCAGCTGCTTTAGATCCATTCATCTACTGTAACCTT 807
QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240
DB CTGCCAAATCACTCACTGACAGATTAATACATTTTACATGCTCATTTGACATCTCTCCCTGC 867
QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
DB ACGACACAGTTGATGATTTTGAATTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
DB GCTGTTTTTTTGAAGTTCTTACATGCAACATCTGTTATGATCATCTGATGAGCTAC 987
QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
DB TTACAAAACAAATTTTCGAGAGCAACAGTACAGTTCTCTAGACAGTTGTTTCTCTCATC 1047
QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
DB ACTGGAAGGAGAGATTTTCATGAGCAGTTTGTAGTTTCAAGAACCCAGAAATGTTTCAGGCT 1107

QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
DB GACCCAGAGATTATACAGCTTCTTGTACATGGGAAGACCTCGAGTCGTTTATGAT 1167
QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
DB ACCATGATTGAGAGTTTGCAGTTTGTACAGCAGTTTGGATGGAGAGCCAAACCAAG 1227
QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
DB CATGAATTTTGCACAGATGGCTATCAGACTTGGTAACATA 1269

RESULT 3

US-10-295-027-1047
; Sequence 1047, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-01250005
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1047
; LENGTH: 3340
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-1047

Alignment Scores:

Pred. No.: 1,52e-217 Length: 3340
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 17 Gaps: 0

US-09-983-000a-2 (1-374) x US-10-295-027-1047 (1-3340)

QY 1 MetArgileLeuLysArgPheLeuAlaCysIleGlnLeuLysCysValCysArgLeuAsp 20
|||||

Db 1 ATCGAATCTTAAGAGCTTTCTCGCTTGCAATCAGCTCCTCTGTGTTTGGCGCTGGAT 60
Qy 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
Db 61 TGGGCTAATGGATACTACAGACAACAGAGAAAATCTGTTGAAGAGATTGGCTGGTCTCTAT 120
Qy 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
Db 121 ACAGAGCAGCTGAATCAAAAATTTGGGGAAGAAATATCCAAATGTAATAGGCCAAAA 180
Qy 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
Db 181 CAATCTCTATCAATATTGATGAAGATCTTACACAGTAATATGTAATCTTAAGAACTT 240
Qy 81 LysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db 241 AAATTTACAGGTTGGGATAAAACATCATTTGGAAAAACATTCATTCATAACACTGGGAAA 300
Qy 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
Db 301 ACAGTGGAAATTAATCTCAATAGTACTACCGTGTACGCGAGGAGTTTCAGAAATGGTG 360
Qy 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
Db 361 TTTAAAGCAAGCAAGATAAATTTTCACTGGGAAAATGCAATATGTCATCTGATGATCA 420
Qy 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 421 GAGCATAGTTTAGAAGGAGCAAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATGCG 480
Qy 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 481 GACCGATTTTCAAGTTTGGAGAGAGCACTCAAGGAAAGGAGTTTAAGAGCTTTATACC 540
Qy 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleAspGlyVal 200
Db 541 ATTTTGTGTTGAGTTGGACAGAGAAAATTTGGATTTCAAGCGATTTATTGATGAGTC 600
Qy 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
Db 601 GAAAGTGTAGTGTGTTTGGGAAGCAGGCTGCTTTAGATCCATTCATCATCTGTAACCTT 660
Qy 221 LeuProAsnSerThrAspLysTyrTyrIleTyrArgGlySerLeuThrSerProProCys 240
Db 661 CTGCCAACTCACTGACAGATTAATACATTTACAAATGCTCATTCAGATCTCCTCCCTGC 720
Qy 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 721 ACACACACAGTTGACTGATTTGTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 780
Qy 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 781 GCTTTTGTGTAAGTCTTACAAATGCAACAATCTGTTATGTCATGCTGATGAGCTAC 840
Qy 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 841 TTACAAAACAATTTTCGAGAGCAACGACAGTACTCTACAGAGTGTCTTCTTCATC 900
Qy 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 901 ACTGAAAGGAAGAGATTCATGAGCAGTTTGTAGTTTCAAGAACAGAAAAATGTTCCAGCT 960
Qy 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
Db 961 GACCCAGAGAAATTAACAGGCTTCTTGTATCATGGGAAAGACCTCGAGTGTATTATGAT 1020
Qy 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 1021 ACCATGATTGAGAGTTTTCGAGTTTGTACACGAGTTTGGATGGAGAGCAACAAACCAAG 1080
Qy 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db 1081 CATGAATTTTTCACAGATGGCTATCAAGACTTGGGTGCTATT 1122

RESULT 4
US-10-295-027-1049
; Sequence 1049, Application US/10295027
; Publication No. US2003023250A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1049
; LENGTH: 4220
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-1049

Alignment Scores:
Pred. No.: 2,18e-217 Length: 4220
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 17 Gaps: 0

US-09-983-000A-2 (1-374) x US-10-295-027-1049 (1-4220)

Qy 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLysValCysArgLeuAsp 20
Db 1 ATGGGAATCTTAAGAGCTTTCTCGCTTGCAATCAGCTCCTCTGTGTTTGGCGCTGGAT 60
Qy 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
Db 61 TGGGCTAATGGATACTACAGACAACAGAGAAAATCTGTTGAAGAGATTGGCTGGTCTCTAT 120
Qy 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
Db 121 ACAGAGCAGCTGAATCAAAAATTTGGGGAAGAAATATCCAAATGTAATAGGCCAAAA 180
Qy 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80

Db 181 CAATCTCTTCAATATTTGATGAAGATCTTACACAGTAAATGTGAATCTTAAGAAACTT 240
Qy 81 LysPheGlnGlyTTPAspLysThrSerLeuGluAenThrPheIleHisenThrGlyLys 100
Db 241 AAATTTTCAGGTTGGGATTAACATCATTTGGAAACACATTCATCAACACTGGGAAA 300
Qy 101 ThrValGluIleAenLeuThrAenAspTyrArgValSerGlyGlyValSerGluMetVal 120
Db 301 ACAGTGGAAATTAATCTCACTAATGACTACCGTGTGCGGAGGAGTTTCAGAAATGGTG 360
Qy 121 PheLysAlaSerLysIleThrPheHisThrGlyLysCysAenMetSerSerAspGlySer 140
Db 361 TTTAAAGCAAGCAAGATACCTTTCACTGGGAAATTCGAATATGATCTCATCTGATGCATCA 420
Qy 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 421 GAGCATAGTTTAAAGGACAAAATTTCCACTTGGATGCAAAATCTACTGCTTTGATGCG 480
Qy 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 481 GACCGATTTTCAAGTTTGGAGAGCAGTCAAGGAAAGGAACTTAAGAGCTTTATCC 540
Qy 181 IleLeuPheGluValGlyThrGluGluAenLeuAspPheLysAlaIleAspGlyVal 200
Db 541 ATTTTGTGTTGAGTTGGGACAGACAGAAATTTGGGATTTCAAAGCGATTAATTGATGAGTC 600
Qy 201 GluSerValSerArgPheGlyGlyGlnAlaLeuAspPropheIleLeuAenLeu 220
Db 601 GAAAGTGTAGTCGTTTGGGAAGCAGGCTGCTTTAGATGCCAATCTATCTGTGGAACCTT 660
Qy 221 LeuProAenSerThrAspLysTyrIleTyrAenGlySerLeuThrSerProCys 240
Db 661 CTGCCAAACTCACTGACAGATTAATACATTTACATGCTCATTTGACATCTCTCCCTGC 720
Qy 241 ThrAspThrValAspTTPilValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 721 ACAGACACAGTTGACTGGATGTTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 780
Qy 261 AlaValPheGlyValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 781 GCTGTTTTTGTGAAGTTCTTTACAAATGCAACAACTCTGTTATGTCTATGCTGATGACTAC 840
Qy 281 LeuGlnAenAspPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 841 TTACAAACAAATTTTCGAGACCAACAGTACAAGTTCTCTAGACAGTGTCTTCTCATAC 900
Qy 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAenValGlnAla 320
Db 901 ACTGGAAGGAAGAGATTCATGAAGCAGTTTGTAGTTCAAGAACCAAGAAATGTTTCAGGCT 960
Qy 321 AspProGluAenTyrThrSerLeuLeuValThrTTPGluArgProArgValValTyrAsp 340
Db 961 GACCCAGAGAATTAATACAGGCTCTCTGTATCATGGGAAAGACCTCGAGTGTATTATGAT 1020
Qy 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 1021 ACCATGATTGAGAAGTTTGCAGTTTGTACAGCAGTTTGGATGGAGAGACCAAAACCAAG 1080
Qy 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db 1081 CATGAATTTTTCACAGATGGCTATCAAGACTTGGGTGCTATT 1122

RESULT 5

US-10-295-027-1042

; Sequence 1042, Application US/10295037

; Publication No. US2003023250A1

; GENERAL INFORMATION:

; APPLICANT: Afar, Daniel

; APPLICANT: Aziz, Natasha

; APPLICANT: Ginsberg, Wendy M.

; APPLICANT: Gish, Kurt C.

; APPLICANT: Glynn, Richard

; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; REMAINING PRIOR APPLICATION DATA REMOVED - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1042
; LENGTH: 4347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-1042

Alignment Scores:

Pred. No.: 2,28e-217 Length: 4347
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 17 Gaps: 0

US-09-983-000A-2 (1-374) x US-10-295-027-1042 (1-4347)

Qy 1 MetArgIleLeuLysArgPheLeuAlaCysTleGlnLeuLeuCysValCysArgLeuAsp 20
Db 1 ATGCGAATCCTAAGACGTTTCTCGCTTGCAATCAGCTCTCTGTGTTCGCCGCTGAT 60
Qy 21 TrpAlaAsnGlyTyrTyrArgGlnArgLysLeuValGluGluIleGlyTTPSerTyr 40
Db 61 TGGGCTAATGGATCTACTACAGCAACAGAGAAACTTGTTCGAGAGATTCGCTGCTCTAT 120
Qy 41 ThrGlyAlaLeuAenGlnLysAenTTPGlyLysTyrProThrCysAenSerProLys 60
Db 121 ACAGGACACTGATCAAAAAAATTTGGGAAAGAAATATCCACATGTAATAGCCCAAAA 180
Qy 61 GlnSerProIleAenIleAspGluAspLeuThrGlnValAenValAenLeuLysLeu 80
Db 181 CAATCTCTATCAATATTGATGAAGATCTTACACAGTAATATGTAATCTTAAGAACTT 240
Qy 81 LysPheGlnGlyTTPAspLysThrSerLeuGluAenThrPheIleHisenThrGlyLys 100
Db 241 AAATTTTCAGGTTGGGATAAAACATCATTTGGAAAAACACATTTCAATACACACTGGGAAA 300
Qy 101 ThrValGluIleAenLeuThrAenAspTyrArgValSerGlyGlyValSerGluMetVal 120
Db 301 ACAGTGGAAATTAATCTCTCACTAATGACTACCGTGTTCAGCGAGGAGTTTCAGAAATGGTG 360

```
QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
Db 361 TTTAAAGCAAGCAAGATAACTTTTCTACGTGGGAAATGCAATATGTCATCTGATGGATCA 420
QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 421 GAGCATAGTTAGAGGACAAAATTTCCATTGAGATGCAAAATCTACTGCTTTGATGCA 480
QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 481 GACCGATTTCAGTTTGGAGAGCAGTCAAGGAAAGGAGTAAAGAGCTTTATACC 540
QY 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIleAspGlyVal 200
Db 541 ATTATTGTTGAGTTGGGACAGAGAAATTTGGATTTCAAAGCGATTAATTGATGGAGTC 600
QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
Db 601 GAAAGTGTAGTCGTTTGGGAGCAGGCTGCTTTAGATCCATTCATCTGTTGACCTT 660
QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240
Db 661 CTGCCAAATCACTCAAGTATTTACATTTACATGGCTCATTTGACATCTCTCCCTGTC 720
QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 721 ACAGACACAGTTGATGGATGTTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 780
QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 781 GCTGTTTTTGTGAGTTCTTACATGCAACATCTGTTATGTCATGCTGATGGACTAC 840
QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 841 TTACAAAACAAATTTTCGAGAGCAACAGTACAAAGTTCTCTAGACAGGTGTTTTCTCTCATAC 900
QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 901 ACTGAAAGGAGAGATTCATGAGCAGTTGTAGTTGAGAACCGAAATGTTTCAGGCT 960
QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
Db 961 GACCCAGAGATTAATACAGCCTCTCTGTTACATCGGAAAGACCTCGAGTGTGTTATGAT 1020
QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 1021 ACCATGATTGAGAGTTTGCAGTTTGTACACAGCAGTTGGATGGAGAGGACCAACCAAG 1080
QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db 1081 CATGAATTTTTCAGAGATGGCTATCAAGACTTTGGGTGCTATT 1122
```

RESULT 6

```
US-10-295-027-1048
; Sequence 1048, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afaf, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
```

```
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1048
; LENGTH: 4480
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-1048

Alignment Scores:
Pred. No.: 2,39e-217 Length: 4480
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 17 Gaps: 0
```

```
US-09-983-000a-2 (1-374) x US-10-295-027-1048 (1-4480)
QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLysValCysArgLeuAsp 20
Db 1 ATGCGAATCTTAAAGCGTTCTCTCGCTTGCAATTCAGCTCCTCTGTGTTGCGCGCTGGAT 60
QY 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
Db 61 TGGGCTAATGATGATCTACAGACACAGAGAAACTTCTTGAAGAGATGGCTGCTCTAT 120
QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysTyrProThrCysAsnSerProLys 60
Db 121 ACAGGACACTGAATCAAAAATAATTTGGGAAAGAAATATCCACATGTAATAGCCCAAAA 180
QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLeu 80
Db 181 CAATCTCTATCAATATTTGATGAAGATCTTTACACAGATAATGTGAATCTTAAAGAAACTT 240
QY 81 LysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db 241 AAATTCAGGTTGGGATAAACAACATCAATGGAATAACATTCATTAACACCTGGGAAA 300
QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
Db 301 ACAGTGGAAATTAATCTCACTAATGACTACCGTGTGAGGAGGAGTTTCAGAAATGGTG 360
QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
Db 361 TTTAAAGCAAGCAAGATAACTTTTCTACTGGGAAAAATGCAATATGTCATCTGATGGATCA 420
QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 421 GAGCATAGTTAGAGGACAAAATTTCCATTGAGATGCAAAATCTACTGCTTTGATGCA 480
QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 481 GACCGATTTTCAAGTTTGGAGAGCAGTCAAGGAAAGGAGTAAAGAGCTTTATACC 540
```



```
QY 181 ileLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIleAspGlyVal 200
Db 541 ATTTGTTGGAGTTGGGACAGAGAAATTTGGATTTCAAGCGATTAATTAATGAGTGC 600
QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspPheIleLeuLeuAsnLeu 220
Db 601 GAAAGTGTAGTCTGTTTGGGAAGCAGGCTGCTTTAGATCCATTCATCTGTTGAACCTT 660
QY 221 LeuProAsnSerThrAspLysTyrrTleTyrrAsnGlySerLeuThrSerProCys 240
Db 661 CTGCCAAACTCAACTGACCAAGTATACATTTACAATGGTTTCATTTGACATCTCTCCCTGC 720
QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 721 ACAGACACAGTTGACTGGATTTGTTTAAAGATACAGATTAGCATCTCTGAAGCCAGTTG 780
QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrrValMetLeuMetAspTyrr 280
Db 781 GCTGTTTTTTTGTGAAGTTCTTACAATGCAACAATCTGGTTATGTCATGCTGATGACTAC 840
QY 281 LeuGlnAsnAspPheArgGluGlnGlnTyrrLysPheSerArgGlnValPheSerTyrr 300
Db 841 TTACAAAACAATTTTCGAGAGCAACAGTACAAGTTCTCTAGACAGGTGTTTCTCCATAC 900
QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 901 ACTGGAAGAGAGAGATTATCATGACGCTTTGATGTTTCAGACCAAGAAATGTTCAAGCT 960
QY 321 AspProGluAsnTyrrThrSerLeuLeuValThrTrpGluArgProArgValValTyrrAsp 340
Db 961 GACCCAGAGAAATATACACGCTTCTTGTATCATGGGAAAGACCTCGAGTCGTTTATGAT 1020
QY 341 ThrMetIleGluLysPheAlaValLeuTyrrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 1021 ACCATGATTGAGAAGTTTCCAGTTTGTACCAAGCAGTTGATGGAGAGACCAACCAAG 1080
QY 361 HisGluPheLeuThrAspGlyTyrrGlnAspLeuValThrIle 374
Db 1081 CATGAATTTTGCAGATGCTTATCAAGACTTTGGGTGCTATT 1122
RESULT 7
US-10-295-027-375
; Sequence 375, Application US/10295027
; Publication No. US2003023350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
```

```
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 375
; LENGTH: 5343
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-375

Alignment Scores:
Pred. No.: 3,13e-217 Length: 5343
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 17 Gaps: 0

US-09-983-000A-2 (1-374) x US-10-295-027-375 (1-5343)
QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
Db 148 ATGCGAATCCTAAAGCGTTTCTCGCTTGCATTCAGCTCCTCTGTGTTTGGCGCTGGAT 207
QY 21 TrpAlaAsnGlyTyrrTyrrArgGlnGlnArgLysValGluGluIleGlyTrpSerTyrr 40
Db 208 TGGCTAATGATACCTACAGACAACAGAGAAACTTTGTTGAAGAGATTGGCTGGTCTTAT 267
QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrrProThrCysAsnSerProLys 60
Db 268 ACAGGAGCAGTGAATCAAAAAAATTTGGGAAAGAAATATCAACATGTAATAGCCCAAAA 327
QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
Db 328 CAATCTCCTATCAATATTGATGAAGATCTTACACAAAGTAAATGTGAATCTTTAAGAAACTT 387
QY 81 LysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db 388 AAATTTCAAGGTTGGATTAACCATCATTTGAAAAACAACATTCATTCATAACACTGGGAAA 447
QY 101 ThrValGluIleAsnLeuThrAsnAspTyrrArgValSerGlyValSerGluMetVal 120
Db 448 ACAGTGGAAATTAATCTCCTCAATGACTACCTGTCGCGGAGGAGTTTCAAGAAATGGTG 507
QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
Db 508 TTTAAAGCAAGCAAGATAACTTTTCTCCTGGGAAAAATGCAATATGTCATCTGATGATCA 567
QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrrCysPheAspAla 160
Db 568 GAGCATAGTTTAGAAGGACAAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATGCA 627
QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 628 GACCGATTTTCAAGTTTGGAGNAGCAGTCAAGAGAAAGGAGGAGTTTAAGAGCTTTATCC 687
QY 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIleAspGlyVal 200
Db 688 ATTTTGTGGAGTTGGGACAGAGAAAAATTTGGATTTCAAGGCGATTTATTGATGAGTGC 747
QY 201 GluSerValSerArgPheGlyLysGlnAlaAlaLeuAspProPheIleLeuLeuAsnLeu 220
Db 748 GAAAGTGTAGTCTGTTTGGGAAGCAGGCTGCTTTAGATCCATTCATCTGATGACCTT 807
QY 221 LeuProAsnSerThrAspLysTyrrTleTyrrAsnGlySerLeuThrSerProCys 240
```


Db 988 TTACAAACAAATTTTCGAGGACCAAGTAGTCAAGATCTCTAGACAGGTGTTTCCCTCATAC 1047
QY 301 ThrGlyLeuGluLeuHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 1048 ACTGGAAAGGAGAGATTATGAGCAGGTTTGTAGTTCAAGACCAAGAAATGTTGAGGCT 1107
QY 321 AppProGluAsnTyrThrSerLeuLeuValThrTTPGluArgProArgValValTyrAsp 340
Db 1108 GACCCAGAGAAATTATACCAAGCTTCTTGTATACATGGGAAAGACCTCGAGTCGTTTATGAT 1167
QY 341 ThrMetIleGluLeuPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 1168 ACCATGATTGAGAGTTTCAGCTTTTGTACCAAGCAGTTCGATGGAGGACCAACCAAG 1227
QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db 1228 CATGAATTTTTCACAGATGGCTATCAAGACTTGGGTGCTATT 1269

RESULT 9

US-10-295-027-1044
; Sequence 1044, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1044
; LENGTH: 5367
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-1044
Alignment Scores:
Pred. No.: 3,15e-217 Length: 5367
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2

Query Match: 99.39% Indels: 0
DB: 17 Gaps: 0
US-09-983-000A-2 (1-374) x US-10-295-027-1044 (1-5367)
QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
Db 148 ATGCGAATCCTAAAGCGTTTCTCGCTTGCAATTCAGCTCCTCTGTGTGTGCGCCTGGAT 207
QY 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysValGluGluIleGlyTyrSerTyr 40
Db 208 TGGGCTAATGATACCTACAGACCAAGAGAAACTTGTGAGAGAGATTGGCTGGTCCCTAT 267
QY 41 ThrGlyAlaLeuAsnGlnLysAsnTyrGlyLysLysTyrProThrCysAsnSerProLys 60
Db 268 ACAGGAGCACTGAATCAAAAATTTGGGAAAGAAATATCCACATGTAAATAGCCCAAA 327
QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
Db 328 CAATCTCCTATCAATATTTGATGAAGATCTTACACAAAGTAAATGTGAATCTTAAAGAACTT 387
QY 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db 388 AAATTTTCAAGGTTGGGATAAAACATCATTTGGGAAACACATTCATTCATAACACTGGGAAA 447
QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
Db 448 ACAGTGGAAATTAATCTCACATAATGACTACCGTGTACAGGAGGAGTTTCAGAAATGGTG 507
QY 121 PheLysAlaSerLysIleThrPheHisTyrGlyLysCysAsnMetSerSerAspGlySer 140
Db 508 TTTAAAGCAAGCAAGATAACTTTTCTCTGGGAAATATGCAATATGTCTCATCTGATGATCA 567
QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 568 GAGCATAGTTTAGAAGGACAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATGCG 627
QY 161 AspArgPheSerSerPheGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 628 GACCGATTTTCAAGTTTGGAGAGCAGTCAAGAGAAAGGAGAGTTTAAAGAGCTTTATCC 687
QY 181 IleLeuPheGluValGlyThrGluAsnLeuAspPheLysAlaIleLeuAspGlyVal 200
Db 688 ATTTTGTGAGGTTGGACAGAGAAAATTTGGATTTCAAAGCGATTTATTGATGAGTGC 747
QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspPheIleLeuLeuAsnLeu 220
Db 748 GAAAGTGTAGTCTGTTTGGGAAAGCAGCTGCTTTAGATCCATTCATCTGTTGAACCTT 807
QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240
Db 808 CTGCCAAACTCACTGACAGATATTAACATTTACATGGCTCATTTGACATCTCTCCCTGCG 867
QY 241 ThrAspThrValAspTyrIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 868 ACAGACACAGTTGACTGGATGTTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 928 GCTGTTTTTGTGAAAGTTCTTACAAATGCAACAAATCTGCTGTTATGTCATGCTCATGACTAC 987
QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 988 TTACAAAACAAATTTTCGAGAGCAACAGTACAGATTTCTTAGACAGGTGTTTTCTCATAC 1047
QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 1048 ACTGGAAGGAAGAGATTTCATGAAGCAGTTTGTAGTTTCAAGAACCAAGAAATGTTTCAGGCT 1107
QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTTPGluArgProArgValValTyrAsp 340
Db 1108 GACCCAGAGAAATTATACCAAGCTTCTTGTGTATACATGGGAAAGACCTCGAGTGTGTTATGAT 1167

QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrIlys 360
DB 1168 ACCATGATTTGAAAGTTTGCAGTTTGTACCGACAGTTTGGATGGAGAGGACCAACCAAG 1227
QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
DB 1228 CATGAATTTTGGACAGATGGCTATCAGACTTGGTGCTATT 1269

RESULT 10

US-10-295-027-379
; Sequence 379, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-0125000US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 379
; LENGTH: 5481
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-379

Alignment Scores:
Pred. No.: 3,25e-217 Length: 5481
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 17 Gaps: 0

US-09-983-000A-2 (1-374) x US-10-295-027-379 (1-5481)

QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuLysValCysArgLeuAsp 20
DB 148 ATGCGAATCCTAAACGGTTTCCGCTTGCATTCAGCTCCTCTGTGTTGGCGCTGGAT 207
QY 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40

DB 208 TGGCTAATGATGACTACTACACACAGAAAACTTTGTTGAAGAGATTGGCTGCTCTAT 267
QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysTyrProThrCysAsnSerProLys 60
DB 268 ACAGGACACTGAATCAAAAAATTTGGGAAAGAAATATCCAAATGTAATAGCCCAAAA 327
QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
DB 328 CAATCTCTATCAATATTGATGAAGATCTTTACACAAGTAATGTGAATCTTAAGAAACTT 387
QY 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
DB 388 AAATTCAGGGTTGGGATAAACAATCATTTGGAAAAACACATTCATTCATAACCTGGGAAA 447
QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
DB 448 ACAGTGGAAATTAATCTCACTAATGACTACCGTGTGACGGGAGGAGTTTCAGAAATGGT 507
QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
DB 508 TTTAAAGCAAGCAAGATAACTTTTCACTGGGAAAAATGCAATATGTTCATCTGATGGATCA 567
QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
DB 568 GAGCATAGTTTAGNAGGACAAAAATTTCCACTTGAGATGCAATCTACTGCTTTGATGG 627
QY 161 AspArgPheSerPheGluGluAlaValLysGlyLysGlyLysLeuAlaLeuSer 180
DB 628 GACCGATTTTCAAGTTTGTAGGAAGCAGTCAAGGAAAAAGGAAGTTTAAGAGCTTTATCC 687
QY 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleLeuAspGlyVal 200
DB 688 ATTTTGTGTTGAGGTTGGGACAGAAAAATTTGGATTTCAAGCGATATTGATGGAGTC 747
QY 201 GluSerValSerArgPheGlyLysGlnAlaAlaLeuAspProPheIleLeuLeuAsnLeu 220
DB 748 GAAAGTGTAGTGGTTTGGGAAGCAGCTGCTTTAGATCCATTCATCTACTGTTGAACCTT 807
QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240
DB 808 CTGCCAAACTCAACTGACAAAGTATTACATTTACAATGGCTCAATGACATCTCTCTCCCTGC 867
QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
DB 868 ACAGACACAGTTGACTGGATTTGTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
DB 928 GCTGTTTTTTGTGAAAGTTCTTACATGCAACAATCTGGTTATGTCATGCTGATGACTAC 987
QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
DB 988 TTAACAAACAAATTTTCGAGAGCAACAGTACAAGTTCTCTAGACAGGTGTTTTCTCATAC 1047
QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
DB 1048 ACTGAAAGGAAGAGAGATTCAATGAAGCAGTTGTAGTTTCAGAACCGAAAAATGTTCAAG 1107
QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
DB 1108 GACCCAGAGAAATATATACAGCCCTTCTTTGTTACATGGGAAAGACCTCGAGTCGTTTATGAT 1167
QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
DB 1168 ACCATGATTCAGAAAGTTTGCAGTTTGTACCAGCAGTTGGATGGAGAGGACCAACCAAG 1227
QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
DB 1228 CATGAATTTTGGACAGATGGCTATCAAGACTTGGGTGCTATT 1269

RESULT 11

US-09-954-456-1812
; Sequence 1812, Application US/09954456

Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
TITLE OF INVENTION: Sets
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1812
LENGTH: 7941
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-1812

Alignment Scores:
Pred. No.: 5.75e-217 Length: 7941
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 9 Gaps: 0

US-09-983-000A-2 (1-374) x US-09-954-456-1812 (1-7941)

QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLysValCysArgLeuAsp 20
DB 148 ATGCGAATCCTAAAGCGTTTCCTCGCTTCGATTCAGCTCCTCTGTGTTTGGCGCGCTGGAT 207

QY 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40
DB 208 TGGCTTAATGGATCTACAGACACAGAGAACTTGTGTAAGAGATTGGCTGGTCTTAT 267

QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysTyrProThrCysAsnSerProLys 60
DB 268 ACAGGAGCACTGAATCAAAAAAATTGGGGAAGAAATATCCAACTGTAATAGCCCAAAA 327

QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
DB 328 CAATCTCCTTATCAATATGATGAAGATCTTACACAAGTAAATGTGAATCTTAAGAAACTT 387

QY 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHiaAsnThrGlyLys 100
DB 388 AATTTTCAGGTTGGATAAACATCATTTGGAAACACATTCATTCATAACACTGGGAAA 447

QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyLysValSerGluMetVal 120
DB 448 ACAGTGGAATTAATCTCACTAATGACTACCGTGTGTCAGCGGAGGATTTCAAGAAATGGTG 507

QY 121 PheLysAlaSerIysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
DB 508 TTTAAAGCAAGCAAGATACTTTTCACTGGGGAAAAATGCAATATGTCATCTCATGGATCA 567

QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
DB 568 GAGCATAGTTTAGAAGGACAAAAATTTCCACTTGTAGATGCAAAATCTACTGCTTGTATGGG 627

QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysGlyLysLeuArgAlaLeuSer 180
DB 628 GACCGAATTTCAAGTTTGTAGGAGCAGTCAAGAGGAAAGGAGGAGTTAAGAGCTTTATCC 687

QY 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleLeuAspGlyVal 200
DB 688 ATTTTGTGTTGAGGTTGGGACAGAAATAATTTGGATTTCAAAGCGGATTAATTGATGGAGTC 747

QY 201 GluSerValSerArgPheGlyLysGlnAlaAlaLeuAspPropheIleLeuLeuAsnLeu 220
DB 748 GAAAGTGTAGTCGTTTGGGAGCAGCGCTGCTTTAGATCCATTCATCTGTTGAACCTT 807

QY 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProCys 240
DB 808 CTGCCAAACTCAACTGACAGATTAATACATTTACAAATGGCTCAATGACATCTCCTCCCTGC 867

QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
DB 868 ACAGACACAGTTGACTGATGTTGTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927

QY 261 AlaValPheCysGluValLeuThrMetGlnSerGlyTyrValMetLeuMetAspTyr 280
DB 928 GCTGTTTTTGTGAAAGTTCTTACAAATGCAACAACTGTTATGTCATGCTGATGACTAC 987

QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
DB 988 TTACAAAACAATTTTCGAGAGCAACAGTACAAAGTTCTCTAGACAGGTGTTTTCCTCATAC 1047

QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
DB 1048 ACTGGAAGGAGGAGATTTCATGAGCAGTTTGTAGTCAGAACCCAGAAATGTTCCAGGCT 1107

QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
DB 1108 GACCCAGAGAAATTAACAGCGCTTCTTGTACATGGGAAAGACCTCGAGTCTGTTATGAT 1167

QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
DB 1168 ACCATGATTGAGAAAGTTTGCAAGTTTGTACCAGCAGTTGGATGGAGAGGACCAACCAAG 1227

QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
DB 1228 CATGAATTTTGGACAGATGGCTATCAAGACTTGGGTGCTATT 1269

RESULT 12
US-09-816-703A-1
Sequence 1, Application US/09816703A
Publication No. US20020146370A1
GENERAL INFORMATION:
APPLICANT: Melcher, Thorsten
APPLICANT: Mueller, Sabine
APPLICANT: Chin, Daniel
TITLE OF INVENTION: Use of Protein Tyrosine Phosphatase Zeta as a Biomolecular Target
FILE REFERENCE: 262/235 AGY
CURRENT APPLICATION NUMBER: US/09/816,703A
CURRENT FILING DATE: 2001-03-23
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 7941
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (148) .. (7092)
US-09-816-703A-1

Alignment Scores:

Pred. No.: 5,75e-217 Length: 7941
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 9 Gaps: 0

US-09-983-000A-2 (1-374) x US-09-816-703A-1 (1-7941)

Qy 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
Db 148 ATGCGAATCCTAAAGCGTTTCTCGCTTGCAATTCAGCTCCTCTGTGTTTGGCGCTGGAT 207
Qy 21 TrpAlaAsnGlyTyrArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
Db 208 TGGGCTAATGGATACTACAGACAACAGAGAAAACCTTGTGTGAAGAGATTGGCTGGCTCTAT 267
Qy 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
Db 268 ACAGGAGCAGCTGAATCAAAAATAATGGGGAAGAAATATCCAAATGTAATAGCCCAAAA 327
Qy 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
Db 328 CAATCTCTATCAATATTGATGAAGATCTTACACAGTAAATGTGAATCTTAAAGAACTT 387
Qy 81 LysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db 388 AAATTTTCAGGGTGGGATAAAACATCATTTGGAAAACACATTCATTCATAAACACTGGGAAA 447
Qy 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyValSerGluMetVal 120
Db 448 ACAGTGGAAATTAATCTCACAATACTACCTGTCGCGGAGGATTTTCAGAAATGGTG 507
Qy 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
Db 508 TTTAAAGCAAGCAAGATAAATTTTCACTGGGAAAATGCAATATGTCATCTGATGATCA 567
Qy 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 568 GAGCATAGTTTAGAAGGACAAAATTTCCACTGAGATGCAAAATCTACTGCTTTGATGCG 627
Qy 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 628 GACCGATTTTCAAGTTTGGAGAGCAGTCAAGGAAAGGAGTTTAAGACTTTATCC 687
Qy 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleAspGlyVal 200
Db 688 ATTTATGTTTGAGTTGGGACAGAGAAAATTTGGATTTCAAAGCGATTATTGATGGAGTC 747
Qy 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
Db 748 GAAAGTGTAGTTCGTTTGGGAAGCAGGCTGCTTTAGATCCATTCATCTGTTGAACCTT 807
Qy 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProProCys 240
Db 808 CTGCCAAACTCACTGACAGATATTACATTTACAAATGCTCATTCAGATCTCTCCCTGCG 867
Qy 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 868 ACAGACACAGTTGACTGGATTGTTTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
Qy 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 928 GCTGTTTTTGTGAAGTTCTTACAAATGCAACAATCTGTTTATGTCATGCTGATGAGCTAC 987
Qy 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 988 TTACAAAACAATTTTCGAGAGCAACAGTACAGTTCTCTACAGAGGTTTTTCTCATAC 1047
Qy 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 1048 ACTGGAAAGGAAGAGATTCATGAAGCAGTTTGTAGTTTCAAGAACAGAAAATGTTTCAGGCT 1107

Qy 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
Db 1108 GACCCAGAGAAATATATACCAGCCTTCTTGTGTACATGGGAAAGACCTCGAGTCGTTTATGAT 1167
Qy 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 1168 ACCATGATTGAGAAAGTTTGCAGTTTGTACACAGCATTTGGATGGAGAGGACCAACCAAG 1227
Qy 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db 1228 CATGAATTTTTCAGACAGATGGCTATCAAGACTTGGGTGCTATT 1269

RESULT 13

US-09-983-000A-5
; Sequence 5, Application US/09983000A
; Publication No. US20030118585A1

GENERAL INFORMATION:

; APPLICANT: AGY Therapeutics
; APPLICANT: Melcher, Thorsten
; APPLICANT: Mueller, Sabine
; APPLICANT: Chin, Daniel
; TITLE OF INVENTION: USE OF PROTEIN BIOMOLECULAR TARGETS IN THE TREATMENT AND VISUALIZ
; TITLE OF INVENTION: OF BRAIN TUMORS
; FILE REFERENCE: 263/180 -- PEAgleman -- AGY
; CURRENT APPLICATION NUMBER: US/09/983,000A
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 7941
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (148)..(7092)
; OTHER INFORMATION:
; NAME/KEY: gene
; LOCATION: (1)..(7941)
; OTHER INFORMATION: PTP-zeta
US-09-983-000A-5

Alignment Scores:

Pred. No.: 5,75e-217 Length: 7941
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 10 Gaps: 0

US-09-983-000A-2 (1-374) x US-09-983-000A-5 (1-7941)

Qy 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
Db 148 ATGCGAATCCTAAAGCGTTTCTCGCTTGCAATTCAGCTCCTCTGTGTTTGGCGCTGGAT 207
Qy 21 TrpAlaAsnGlyTyrArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
Db 208 TGGGCTAATGGATACTACAGACAACAGAGAAAACCTTGTGTGAAGAGATTGGCTGGCTCTAT 267
Qy 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
Db 268 ACAGGAGCAGCTGAATCAAAAATAATGGGGAAGAAATATCCAAATGTAATAGCCCAAAA 327
Qy 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
Db 328 CAATCTCTATCAATATTGATGAAGATCTTACACAGTAAATGTGAATCTTAAAGAACTT 387
Qy 81 LysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db 388 AAATTTTCAGGGTGGGATAAAACATCATTTGGAAAACACATTCATTCATAAACACTGGGAAA 447
Qy 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyValSerGluMetVal 120

```
Db 448 ACAGTGGAAATTAATCTCACTAATGATACCGTGTGAGCGGAGGAGTTTCAGAAATGGTG 507
Qy 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
Db 508 TTAAAGCAAGCAAGTAATCTTTCTCTGGGAAATGCAATATGTCATCTCATGTGATCA 567
Qy 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 568 GAGCATAGTTTAGAAGGACAAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATGCG 627
Qy 161 AspArgPheSerSerPheGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 628 GACCGGATTTCAAGTTTTCAGGAGCAGTCAAGAGGAAATGGAAGTTTAGAGCTTTATGCC 687
Qy 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleLeuAspGlyVal 200
Db 688 ATTTGTTTGAGTTGGGACAGAGAAATTTGGATTTCAAGCGAATATTGATGGAGTC 747
Qy 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
Db 748 GAAAGTGTAGTCGTTTGGGAGCAGGCTGCTTTAGATCCATTCATCTGTTGAACCTT 807
Qy 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProCys 240
Db 808 CTGCCAAACTCACTGACCAAGTATTACATTTACAATGGCTCATTCACATCTCTCCCTGC 867
Qy 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 868 ACAGACACAGTTGACTGGATTTGTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
Qy 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 928 GCTGTTTTTGTGAAGTCTTCAATGCAACCAATCTGGTTATGTCATGCTGATGACTAC 987
Qy 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 988 TTACAAAAACAATTTTCGAGAGCAACAGTACAAAGTTCTCTAGACAGGTGTTTTCCTCAT 1047
Qy 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 1048 ACTGGAAGGAGAGAGATTCATGAGCAGTTTGTAGTTCAAGACCAGAAATGTTCAAGCT 1107
Qy 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
Db 1108 GACCCAGAGAAATTATACAGCTTCTTGTTCATGCGGAAAGACCTCGAGTCGTTATGAT 1167
Qy 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 1168 ACCATGATTGAGAAGTTTGCACTTTGTACCAAGCAGTTGGATGGAGAGACCAACCAAG 1227
Qy 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db 1228 CATGAATTTTTCACAGATGGCTATCAAGACTTTGGGTGCTATT 1269
```

RESULT 14

```
US-10-295-027-381
; Sequence 381, Application US/10295027
; Publication No. US2003023250A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
```

```
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 381
; LENGTH: 7941
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-381
```

```
Alignment Scores:
Pred. No.: 5,75e-217 Length: 7941
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 17 Gaps: 0

US-09-983-000A-2 (1-374) x US-10-295-027-381 (1-7941)
```

```
Qy 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLeuCysValCysArgLeuAsp 20
Db 148 ATGCGAATCCCTAAAGCGTTTCCTCGCTTGCAATCAGCTCCTCTGTGTGTGCGCTGGAT 207
Qy 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluIleGlyTyrSerTyr 40
Db 208 TGGGCTAATGGATACTACAGACAAAGAGAAACTTGTGGAAGAGATTGGCTGCTCTAT 267
Qy 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
Db 268 ACAGAGCACTGATCAAAAAAATTTGGGAAAGAAATATCCACATGTATAGGCCAANA 327
Qy 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
Db 328 CAATCTCTATCAATATTGATGAAGATCTTTACACAGTAAATGTGAATCTTTAAGAAACTT 387
Qy 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db 388 AAATTTTCAGGGTTGGGATAAAACATCATTTGGAAACACATTCATTCATAACACTGGGAAA 447
Qy 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
Db 448 ACAGTGGAAATTAATCTCACTAATGACTACCGTGTGCGGAGGAGGTTTCAGAAATGGTG 507
Qy 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
Db 508 TTTAAAGCAAGCAAGATAACTTTTTCACCTGGGAAATATGCAATATGTCATCTCATGTGATCA 567
Qy 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
Db 568 GAGCATAGTTTAGAAGGACAAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATGCG 627
```



```
Qy 161 AspArgPheSerSerPheGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 628 GACCGATTTTCAAGTTTGTGAGAACAGTCAAGAAAGAGGAGTTAAGAGCTTTATCC 687
Qy 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleAspGlyVal 200
Db 688 ATTTTGTGTGAGTTGGGACAGAGAAATTTGGATTTCAAAGCGATTTATGATGGATC 747
Qy 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspPheIleLeuLeuAsnLeu 220
Db 748 GAAAGTGTAGTCGTTTGGGAAGCAGCGCTGTTTAGATCCATTCATCATCTGTTGAACCTT 807
Qy 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProCys 240
Db 808 CTGCCAAACTCAACTGACAGATTTACATTTTACAAATGGCTCATTCACATCTCCCTCCGTC 867
Qy 241 ThrAspThrValAspTyrIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 868 ACAGACACAGTTGACTGGATGTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
Qy 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 928 GCTGTTTTTGTGAAGTCTTACAAATGCAACAATCTGTTATGTATGTCATGCTGATGACTAC 987
Qy 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 988 TTACAAAACAATTTTCGAGAGCAACAGTACAGTTCTCTAGACAGGTGTTTCTCTCATAC 1047
Qy 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 1048 ACTGGAAGGAGAGATTCATGACAGCTTTGTAGTTCAGAACCCAGAAATGTTCCAGCT 1107
Qy 321 AspProGluAsnTyrThrSerLeuLeuValThrTyrGluArgProArgValValTyrAsp 340
Db 1108 GACCCAGAGAAATTATACAGCCCTTCTGTTATCATGGGAAAGACCTCGAGTCTGTTATGAT 1167
Qy 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 1168 ACCATGATTGAGAGTTTGCAGTTTGTACACAGCAGTTGGATGGAGAGGACCAACCAAG 1227
Qy 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db 1228 CATGAATTTTGACAGATGGCTATCAAGACTTGGGTGCTATT 1269

RESULT 15
US-10-843-641A-4839
; Sequence 4839, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843, 641A
; CURRENT FILING DATE: 2004-05-12
; PRIORITY FILING DATE: US/09/873,367
; PRIORITY FILING DATE: 2001-06-05
; PRIORITY FILING DATE: US/09/954,531
; PRIORITY FILING DATE: 2001-09-18
; PRIORITY FILING DATE: US/09/954,456
; PRIORITY FILING DATE: 2001-09-25
; PRIORITY FILING DATE: US/09/962,436
; PRIORITY FILING DATE: 2001-09-25
; PRIORITY FILING DATE: US/09/962,832
; PRIORITY FILING DATE: 2001-09-25
; PRIORITY FILING DATE: US/09/964,824
; PRIORITY FILING DATE: 2001-09-27
; PRIORITY FILING DATE: US/09/967,768
; PRIORITY FILING DATE: 2001-09-28
; PRIORITY FILING DATE: US/09/968,007
; PRIORITY FILING DATE: 2001-10-02
; PRIORITY FILING DATE: US/09/969,347
; PRIORITY FILING DATE: 2001-10-02
```

```
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4839
; LENGTH: 7941
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-4839

Alignment Scores:
Pred. No.: 5,75e-217 Length: 7941
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 21 Gaps: 0
```

US-09-983-000A-2 (1-374) x US-10-843-641A-4839 (1-7941)

```
Qy 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuCysValCysArgLeuAsp 20
Db 148 ATGCGAATCTTAAAGCGTTTCTCGCTTGCATTGAGCTCTCTGTTGCGGCTGGAT 207
Qy 21 TrpAlaAsnGlyTyrTyrArgGlnGlnArgLysLeuValGluGluIleGlyTyrSerTyr 40
Db 208 TGGGCTTAATGGTACTACACACACAGAGAAAATTTGTTGAAGAGATTGGCTGGTCTCTAT 267
Qy 41 ThrGlyAlaLeuAsnGlnLysAsnTyrGlyLysLysTyrProThrCysAsnSerProLys 60
Db 268 ACAGGACACTGATTCAAAAAATTTGGGAAAGAAATATCCACATGTATATAGCCCAAAA 327
Qy 61 GluSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLeu 80
Db 328 CAATCTCTATCAATATTGATGAAGATCTTACACAGTAAATGTGAATCTTAAAGAAACTT 387
Qy 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
Db 388 AAATTTTCAGGGTTGGGATTAACAATCATTTGGAACACATTCATTCATAACACACACAC 447
Qy 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
Db 448 ACAGTGGAAATTAATCTCATTATGACTACCGTGTACGCGGAGGAGTTTCAGAAATGGTG 507
Qy 121 PheLysAlaSerLysIleThrPheHisTyrGlyLysCysAsnMetSerSerAspGlySer 140
Db 508 TTTAAAGCAAGCAAGATAACTTTTCTACTGGGAAAAATGCAATATGTCTCATCTGTATGGATCA 567
Qy 141 GluHisSerLeuGluGlyGlnLysPheProIleGluMetGlnIleTyrCysPheAspAla 160
Db 568 GAGCATAGTTTAGAAGACAAAAATTTCCACTTGAGATGCAAAATCTACTGCTTTGATGCG 627
Qy 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 628 GACCGATTTTCAAGTTTGTGAGAACAGTCAAGGAAAAAGGAGATTAAAGAGCTTTATCC 687
Qy 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleAspGlyVal 200
Db 688 ATTTTGTGTGAGTTGGGACAGAGAAATTTGGATTTCAAAGCGATTTATGATGGATC 747
Qy 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspPheIleLeuLeuAsnLeu 220
Db 748 GAAAGTGTAGTCGTTTGGGAAGCAGCGCTGTTTAGATCCATTCATTCATCTGTTGAACCTT 807
Qy 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProCys 240
Db 808 CTGCCAAACTCAACTGACAGATTTACATTTTACAAATGGCTCATTCACATCTCCCTCCGTC 867
Qy 241 ThrAspThrValAspTyrIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 868 ACAGACACAGTTGACTGGATGTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 927
```


QY	261	AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr	280
Db	928	GCATGTTTTTTGTGAAGTCTTACAAATGCAACAATCTGGTTATGTCTATGCTGATGGACTAC	987
QY	281	LeuGlnAsnAsnPheArgGluGlnTyrLysPheSerArgGlnValPheSerSerTyr	300
Db	988	TTACAAACAACATTTTCGAGAGCAACAGTACAAGTTCTCTAGACAGGTGTTTTCTCATAC	1047
QY	301	ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla	320
Db	1048	ACTGGAAAGCAAGAGATTCTAGACAGTTTGTAGTTTCAGAACCAAGAAATGTTCAGGCT	1107
QY	321	AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp	340
Db	1108	GACCCAGAGAAATTATACCAGCCTTCTTGTTCATATGGGAAAGACCTCGAGTCGTATTATGAT	1167
QY	341	ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys	360
Db	1168	ACCATGATTGAGAGTTTTCAGTTTTGTACCAAGCAGTTGGATGGAGAGGACCAACCAAG	1227
QY	361	HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle	374
Db	1228	CATGAATTTTTCACAGATGGCTATCAGACTTTGGGTGCTATT	1269

Search completed: August 30, 2005, 12:17:31
Job time : 723 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 30, 2005, 08:31:56 ; Search time 207 Seconds
(without alignments)
2956.365 Million cell updates/sec

Title: US-09-983-000A-2

Perfect score: 1980

Sequence: 1 MRILKRFLACQLLCVCRLD.....GEDQTKHEFLTDGYQLVTI 374

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgm2_1/USPTO_spool_p/US0983000/runat_29082005_155924_14685/app_query.fasta_1.519
-DB=Issued Patents NA -QW=fastcap -SUFFIX=rn1 -MINWATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNIT9=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTWT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0983000 @CGN 1.1.105 @runat_29082005_155924_14685 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgm2_6/ptodata/1/ina/5A COMB.seq:
2: /cgm2_6/ptodata/1/ina/5B COMB.seq:
3: /cgm2_6/ptodata/1/ina/6A COMB.seq:
4: /cgm2_6/ptodata/1/ina/6B COMB.seq:
5: /cgm2_6/ptodata/1/ina/PTUS COMB.seq:
6: /cgm2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1968	99.4	6924	1	US-08-015-973-2
2	1968	99.4	6924	2	US-08-448-164-2
3	1968	99.4	7941	3	US-09-815-703A-1
4	648.5	32.8	4338	1	US-08-015-986A-1
5	648.5	32.8	4338	2	US-08-446-363-1
6	412	20.8	813	1	US-08-276-919-9
7	412	20.8	813	1	US-08-776-088-12
8	412	20.8	813	5	PCT-US95-09145A-12
9	412	20.8	822	1	US-08-776-088-14
10	412	20.8	822	5	PCT-US95-09145A-14
11	412	20.8	986	1	US-08-276-919-3
12	412	20.8	986	1	US-08-276-919-12

13	412	20.8	986	1	US-08-776-088-3	Sequence 3, Appli
14	412	20.8	986	1	US-08-776-088-17	Sequence 17, Appli
15	412	20.8	986	5	PCT-US95-09145A-3	Sequence 3, Appli
16	412	20.8	986	5	PCT-US95-09145A-17	Sequence 17, Appli
17	412	20.8	1104	1	US-08-276-919-1	Sequence 1, Appli
18	412	20.8	1104	1	US-08-776-088-1	Sequence 1, Appli
19	412	20.8	1104	3	US-09-325-320-1	Sequence 1, Appli
20	412	20.8	1104	3	US-09-585-109-1	Sequence 1, Appli
21	412	20.8	1104	5	PCT-US95-09145A-1	Sequence 1, Appli
22	412	20.8	2134	1	US-08-776-088-5	Sequence 5, Appli
23	412	20.8	2134	5	PCT-US95-09145A-5	Sequence 5, Appli
24	412	20.8	2774	4	US-09-949-016-4029	Sequence 4029, Ap
25	412	20.8	2775	4	US-09-949-016-172	Sequence 172, App
26	400	20.2	2679	1	US-08-479-328-1	Sequence 1, Appli
27	400	20.2	2679	1	US-08-761-119-1	Sequence 1, Appli
28	400	20.2	2679	2	US-08-668-128B-1	Sequence 1, Appli
29	400	20.2	2679	2	US-08-905-445-1	Sequence 1, Appli
30	400	20.2	2679	3	US-08-959-625-1	Sequence 1, Appli
31	400	20.2	2679	3	US-09-008-466-1	Sequence 1, Appli
32	400	20.2	2679	3	US-08-580-980A-1	Sequence 1, Appli
33	400	20.2	2679	3	US-09-053-453-1	Sequence 1, Appli
34	400	20.2	2679	3	US-08-644-116A-1	Sequence 1, Appli
35	400	20.2	2679	4	US-03-437-858-1	Sequence 1, Appli
36	396.5	20.0	1190	2	US-08-668-128B-4	Sequence 4, Appli
37	396.5	20.0	1190	2	US-08-905-445-4	Sequence 4, Appli
38	396.5	20.0	1190	3	US-08-959-625-4	Sequence 4, Appli
39	396.5	20.0	1190	3	US-09-008-466-4	Sequence 4, Appli
40	396.5	20.0	1190	3	US-08-580-980A-4	Sequence 4, Appli
41	396.5	20.0	1190	3	US-09-053-453-4	Sequence 4, Appli
42	396.5	20.0	1190	3	US-08-644-116A-4	Sequence 4, Appli
43	396.5	20.0	1190	4	US-09-437-858-4	Sequence 4, Appli
44	312.5	15.8	1428	4	US-09-949-016-728	Sequence 728, App
45	304.5	15.4	1399	3	US-08-335-469-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-015-973-2
; Sequence 2, Application US/08015973
; Patent No. 5604094
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-BETA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/015,973
; FILING DATE: 10-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212)869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6924 base pairs

;; TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6924
US-08-015-973-2

Alignment Scores:

Pred. No.: 4.05e-232 Length: 6924
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 1 Gaps: 0

US-09-983-000A-2 (1-374) x US-08-015-973-2 (1-6924)

QY 1 MetArgIleLeuLysArgPheLeuAlaCysIleGlnLeuLysValCysArgLeuAsp 20
DB 1 ATGGGAATCCTAAAGCGTTTCCTCGCTTGCATTGAGCTCCTCTGTGTTGCGCGCTGGAT 60
QY 21 TrpAlaAenGlyTyrTrpArgGlnGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
DB 61 TGGGCTAATGATACTACAGACACAGAGAAAATCTGTGTGAAGATTTGGCTGCTCTAT 120
QY 41 ThrGlyAlaLeuAenGlnLysAenTrpGlyLysLysTyrProThrCysAenSerProLys 60
DB 121 ACAGGAGCACTGAATCAAAAATTTGGGAAAGAAATATCCAAATGTAATAGCCCAAAA 180
QY 61 GlnSerProIleAenIleAspGluAspLeuThrGlnValAenValAenLeuLysLysLeu 80
DB 181 CAATCTCCTATCAATATTATGATGAAGATCTTACACAGTAATATGTAATCTTTAAGAAACTT 240
QY 81 LysPheGlnGlyTrpAspLysThrSerLeuGluAenThrPheIleHisAenThrGlyLys 100
DB 241 AAATTCAGGGTGGGTAAACATCATTTGGAACACATTCATTCATCAACACTGGGAAA 300
QY 101 ThrValGluIleAenLeuThrAsnAspTyrArgValSerGlyLysValSerGluMetVal 120
DB 301 ACAGTGGAAATTAATCTCACTAATGACTACCGTGTGACGGGAGGATTTACAGAAATGGT 360
QY 121 PheLysAlaSerLysIleThrPheHIStrpGlyLysCysAenMetSerSerAspGlySer 140
DB 361 TTTAAAGCAAGACAGATAACTTTTTCACGTGGGAAATGCAATATGTCATCTGATGGATCA 420
QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
DB 421 GAGCACAGTTTAGAAGGACAAATTTCCATTGAGATGCAAAATCTACTGCTTTGATGCA 480
QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
DB 481 GACCGATTTTCAAGTTTGGAGAGCAGTCAAGGAAAGGAAAGTAAAGAGCTTTATATCC 540
QY 181 IleLeuPheGluValGlyThrGluGluAenLeuAspPheLysAlaIleLeuAspGlyVal 200
DB 541 ATTTGTTTGGAGTTGGGACAGAGAAATTTGGAATTTCAAGGCGATTATTGATGGAGTC 600
QY 201 GluSerValSerArgPheGlyLysGlnAlaAlaLeuAspProPheIleLeuLeuAenLeu 220
DB 601 GAAAGTGTAGTCTGTTTGGAGACAGCGCTCTCTAGATCCATTCATCTACTGTTGACCTT 660
QY 221 LeuProAenSerThrAspLysTyrTyrIleTyrAenGlySerLeuThrSerProProCys 240
DB 661 CTGCCAAACTCAACTGACCAAGTATTACATTTTACAAATGGCTCATTTGACATCTCTCCCTGC 720
QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
DB 721 ACAGACACAGTTGATGGATTTGTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 780
QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280

DB 781 GCTGTTTTTGTGAAGTTCTTACATGCAACAATCTGGTTATGTCATGCTGATGAGCTAC 840
QY 281 LeuGlnAenAenPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
DB 841 TTACAAAACAATTTTCGAGAGCAACAGTCAAGTTCTCTAGACAGGTGTTTTCTCTCATAC 900
QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAenValGlnAla 320
DB 901 ACTGGAAGGAAGAGATTCAATGAAGCAGTTTGTAGTTTCAGAACCAAAAATGTTGAGCT 960
QY 321 AspProGluAenTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
DB 961 GACCCAGAGATTTATACCAAGCTTCTTGTATACATGGGAAAGACCTCGAGTCTGTTATGAT 1020
QY 341 ThrMetIleGlyLysPheAlaValLeuTyrGlnGlnLeuAspGlyLysGlnThrLys 360
DB 1021 ACCATGATTGAGAAGTTTGCAGTTTGTACCAAGCAGTTGGATGGAGAGGACCAACCAAG 1080
QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
DB 1081 CATGAATTTTTCAGACATGGCTATCAGACTTGGGTGCTATT 1122
RESULT 2
US-08-448-164-2
; Sequence 2, Application US/08448164
; Patent No. 5925536
; GENERAL INFORMATION:
; APPLICANT: Schliesinger, Joseph
; TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-BETA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,164
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/015,973
; FILING DATE: 10-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212)869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6924 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..6924
US-08-448-164-2
Alignment Scores: 4.05e-232 Length: 6924
Pred. No.: 1968.00 Matches: 372
Score:

Percent Similarity: 99.47% Conservatives: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 2 Gaps: 0

US-09-983-000A-2 (1-374) x US-08-448-164-2 (1-6924)

```
QY 1 MetArgileLeuLysArgPheLeuAlaCysIleGlnLeuLysValCysArgLeuLeu 20
DB 1 ATGCGAATCTTAAAGGTTTCTCGCTTGCATTCAGCTCTCTGTGTTTGGCGCTGGAT 60
QY 21 TrpAlaAsnGlyTyrTrpArgGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
DB 61 TGGGCTTAATGGATACACAGACAGAGAAATCTGTTGAAGAGATTGGCTGGCTCTAT 120
QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
DB 121 ACAGGACGACTGATCAATCAAAAATTTGGGAAAGAAATATCCAACTGTAATAGCCAAA 180
QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
DB 181 CAATCTCTATCAATATTCATGAAGATCTTACACAGTAAATGTGAATCTTAAGAACTT 240
QY 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
DB 241 AAATTTTCAAGGTTGGGATAAACATCATTTGGAAAAACACATTCATTCATAAACA 300
QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
DB 301 ACAGTGGAAATTAATCTCACTAATGACTACCTGCTGAGCGGAGGAGTTTCAGAAATGGTG 360
QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
DB 361 TTTAAAGCAAGCAAGATACTTTTCACTGGGGAATATGCAATATGTCATCTGATGGATCA 420
QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
DB 421 GAGCAGCTTTAAGAGGACAAAAATTTCCACTTGAGATGCAATCTACTGCTTTGATGCA 480
QY 161 AspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
DB 481 GACCCATTTTCAAGTTTGGAGAGCAGTCAAGGAAAGGAAAGTTAAGAGCTTTATACC 540
QY 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleAspGlyVal 200
DB 541 ATTTGTTTGGAGTTGGACAGACAGAAATTTGGATTTCAAGCGATTAATGATGGATGC 600
QY 201 GluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeuLeuAsnLeu 220
DB 601 GAAAGTGTAGTCGTTTGGGAGCAGGCTGCTTTAGATCCATTCATCTGTTGAACCTT 660
QY 221 LeuProAsnSerThrAspLysTyrIleTyrAsnGlySerLeuThrSerProCys 240
DB 661 CTGCCAACTCACTGACAGATTAATACATTTTACAATGGCTCAATGACATCTCTCCCTGCG 720
QY 241 ThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
DB 721 ACAGACACAGTTGACTGGATTTTAAAGATACAGTTAGCATCTCTGAAAGCCAGTTG 780
QY 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
DB 781 GCTGTTTTTGTGAAGTCTTACAATGCAACAATCTGGTTATGTCATGCTGATGACTAC 840
QY 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
DB 841 TTACAAAACAAATTTTCGAGAGCAACAGTACAGTTCTCTAGACAGGTGTTTCTCATAC 900
QY 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
DB 901 ACTGGAAGGAGAGATTCATGACAGTTTGTAGTTTCAAGAACCAAGAAATGTTTCAGGCT 960
QY 321 AspProGluAsnTyrThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
```

```
DB 961 GACCCAGAGAAATTATACAGGCTTCTTGTTACATGGGAAAGACCTCGAGTCGTTTATGAT 1020
QY 341 ThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
DB 1021 ACCATGATTGAGAGTTTGCAGTTTGTACACGACAGTTTGGATGGAGAGCCAAACCAAG 1080
QY 361 HisGluPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
DB 1081 CATGAATTTTGGACAGATGGCTATCAAGACTTGGGTGCTATT 1122
```

RESULT 3

US-09-816-703A-1
; Sequence 1, Application US/09816703A
; Patent No. 6455026
; GENERAL INFORMATION:
; APPLICANT: Melcher, Thorsten
; APPLICANT: Mueller, Sabine
; APPLICANT: Chin, Daniel
; TITLE OF INVENTION: Use of Protein Tyrosine Phosphatase Zeta as a Biomolecular Target
; FILE REFERENCE: 262/235 AGY
; CURRENT APPLICATION NUMBER: US/09/816,703A
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 7941
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (148)..(7092)
US-09-816-703A-1

Alignment Scores:

Pred. No.: 5,04e-232 Length: 7941
Score: 1968.00 Matches: 372
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 3 Gaps: 0

US-09-983-000A-2 (1-374) x US-09-816-703A-1 (1-7941)

```
QY 1 MetArgileLeuLysArgPheLeuAlaCysIleGlnLeuLysValCysArgLeuLeu 20
DB 148 ATGCGAATCTTAAAGGTTTCTCGCTTGCATTCAGCTCTCTGTGTTTGGCGCTGGAT 207
QY 21 TrpAlaAsnGlyTyrTrpArgGlnArgLysLeuValGluGluIleGlyTrpSerTyr 40
DB 208 TGGGCTTAATGGATACACAGACAGAGAAATCTGTTGAAGAGATTGGCTGGCTCTAT 267
QY 41 ThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsnSerProLys 60
DB 268 ACAGGACGACTGATCAATCAAAAATTTGGGAAAGAAATATCCAACTGTAATAGCCAAA 327
QY 61 GlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeuLysLysLeu 80
DB 328 CAATCTCTATCAATATTCATGAAGATCTTACACAGTAAATGTGAATCTTTAAGAACTT 387
QY 81 LysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsnThrGlyLys 100
DB 388 AAATTTTCAAGGTTGGGATAAACATCATTTGGAAAAACATTCATTCATAAACA 447
QY 101 ThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSerGluMetVal 120
DB 448 ACAGTGGAAATTAATCTCACTAATGACTACCGTGTGAGCGGAGGAGTTTCAGAAATGGTG 507
QY 121 PheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSerAspGlySer 140
DB 508 TTTAAAGCAAGCAAGATACTTTTCACTGGGGAATATGCAATATGTCATCTGATGGATCA 567
QY 141 GluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCysPheAspAla 160
```

```
Db 568 GAGCATAGTTTGAAGGACAAATTTCCATCTGAGATGCAATCTACTGCTTTGATCG 627
Qy 161 AspArgPheSerPheGluGluAlaValLysGlyLysGlyLysLeuArgAlaLeuSer 180
Db 628 GACCGATTTTCAAGTTTGTGAGAGCAGTCAAGGAAAGGAAAGTTAAGAGCTTTATCC 687
Qy 181 IleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleAspGlyVal 200
Db 698 ATTGTGTTGAGTTGGGACAGAGAAATTTGGATTTCAAAGCGATTATTGATGGAGTC 747
Qy 201 GluSerValSerArgPheGlyLysGlnAlaAlaLeuAspProPheIleLeuLeuAsnLeu 220
Db 748 GAAAGTGTAGTCGTTTGGAGAGCAGCTCTCTTAGATCCATTCATCTGTTGACCTT 807
Qy 221 LeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSerProCys 240
Db 808 CTGCCAAACTCACTGACCAAGTATTACATTTACAATGGCTCATTTGACATCTCCTCCCTGC 867
Qy 241 ThrAspThrValAspTyrIleValPheLysAspThrValSerIleSerGluSerGlnLeu 260
Db 868 ACAGACACAGTTGACTGAGATTGTTTAAAGACATACAGTTAGCATCTCTGAAAGCCAGTTG 927
Qy 261 AlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyr 280
Db 928 GCTGTTTTTTGTAAGTTCTTACATGCAATGCAATCTGGTTAATGATGCTGATGACTAC 987
Qy 281 LeuGlnAsnAsnPheArgGluGlnGlnTyrLysPheSerArgGlnValPheSerSerTyr 300
Db 988 TTACAAACAATTTTCGAGAGCAACAGTACAAAGTTCTCTAGACAGGTGTTTTCTCTATAC 1047
Qy 301 ThrGlyLysGluGluIleHisGluAlaValCysSerSerGluProGluAsnValGlnAla 320
Db 1048 ACTGAAAGGAGAGATTCATGAAGCAGTTTGTAGTTTCAGAACAGAAATGTTTCAGCT 1107
Qy 321 AspProGluAsnThrSerLeuLeuValThrTrpGluArgProArgValValTyrAsp 340
Db 1108 GACCAGAGATTTATACAGCTCTCTGTTATCATGGGAAAGACCTCGAGTCTGTTATGAT 1167
Qy 341 ThrIleGluLysPheAlaValLeuTyrGlnGlnLeuAspGlyGluAspGlnThrLys 360
Db 1168 ACCATGATGAGAGTTTGCAGTTTGTACAGCAGTGTGGATGGAGAGACCAACCAAG 1227
Qy 361 HisLeuPheLeuThrAspGlyTyrGlnAspLeuValThrIle 374
Db 1228 CATCAATTTTGGACAGATGGCTATCAAGACTTGGGTGCTATT 1269
```

RESULT 4

```
US-08-015-986A-1
; Sequence 1, Application US/08015986A
; Patent No. 5532123
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-GAMMA
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESS: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/015,986A
; FILING DATE: 10-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 19,872
; REFERENCE/DOCKET NUMBER: 7683-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4338 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4335
; US-08-015-986A-1

Alignment Scores:
Pred. No.: 3,73e-69 Length: 4338
Score: 648.50 Matches: 136
Percent Similarity: 55.31% Conservative: 62
Best Local Similarity: 37.99% Mismatches: 145
Query Match: 32.75% Indels: 15
DB: 1 Gaps: 2

US-09-983-000A-2 (1-374) x US-08-015-986A-1 (1-4338)
Qy 28 GlnGlnArgLysLeuValGluGluIleGlyTyrSerTyrThrGlyAlaLeuAsnGlnLys 47
Db 148 CGCAGGCGCAGGCTTCAGCGACCCGCTACTGGCTTACTCTGGTGGCTATGGTCTGAG 207
Qy 48 AsnTrpGlyLysLysTyrProCysAsnSerProLysGlnSerProIleAsnIleAsp 67
Db 208 CACTGGGTACGCTAGTGTGCTGAGCGGCGCTCACCAGTCTCTCTATTGACATTTTA 267
Qy 68 GluAspLeuThrGlnValAsnValAsnLeuLysLysLysPheGlnGlyTyrAspLys 87
Db 268 GACCAGTATCGCGGTGTGGGAGAAATACAGAACTGCAACTCGATGGCTTCGCAAT 327
Qy 88 ThrSerLeuGluAsnThrPheIleHisAsnThrGlyLysThrValGluIleAsnLeuThr 107
Db 328 GAGTCTTCTAACAAACCTGATGAAACACACAGGAAACAGTCGCCATCTCTCGAA 387
Qy 108 AsnAspTyrArgValSerGlyValSerGluMetValPheLysAlaSerLysIleThr 127
Db 388 GACGACTATTTTGTGCTGAGTGGCTGTCTACTGGCAGATTCAAGCTGAGAGGTGAA 447
Qy 128 PheHisTrpGlyLysCysAsnMetSerSerAspGlySerGluHisSerLeuGluGln 147
Db 448 TTTCACTGGGGCCACAGCAAT--GGCTCAGCGGCTCTGAACACAGCATCAATGGCAGG 504
Qy 148 LysPheProLeuGluMetGlnIleTyrCysPheAspAlaAspArgPheSerSerPheGlu 167
Db 505 AGGTTTCTGTTGAGATGAGATTTCTTTTCAATCCAGATGACTTTGACAGCTTCAA 564
Qy 168 GluAlaValLysGlyLysGlyLysLeuArgAlaLeuSerIleLeuPheGluValGlyThr 187
Db 565 ACCGCAATTTCTGAGAACAGATAATCGAGCGCCATATTTTTCAGTCAGTCCG 624
Qy 188 GluGluAsnLeuAspPheLysAlaIleAspGlyValGluSerValSerArgPheGly 207
Db 625 AGGACAAATTTCTGCATCTGATTCCTATTATCCAGGGTTGAAGGTGCTGATCATGAG 684
Qy 208 LysGlnAlaAlaLeuAspProPheIleLeuLeuAsnLeuProAsnSerThrAspLys 227
Db 685 AAGAGACCTTTCTGGATCTCTTCGCTCCGGACCTCTCGCTGCTGATCCCTGGGAGC 744
Qy 228 TyrTyrIleTyrAsnGlySerLeuThrSerProCysThrAspThrValAspTyrIle 247
Db 745 TATTATCGGTACACAGGTTCTTGGACCAACACCGTGTAGCGAAATAGTGGATGGATA 804
```

QY 248 ValPheLysAspThrValSerIleSerGluSerGlnLeuAlaValPheCysGluValLeu 267
 DB 805 GTCTTCGGGAGACCGTCCCATCTCTTACCATCAGCTTGAGGCTTTTATTCCATCTTC 864
 QY 268 ThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyrIleGlnAsnAsnPheArgGlu 287
 DB 865 ACCACGGACGACGACGACCATGTCAAGTCGGTGAGATCTGAGAAATAACTTTTCGACCA 924
 QY 288 GlnGlnTyrLysPheSerArgGlnValPheSerSerTyrThrGlyLysGluGluIleHis 307
 DB 925 CAGCAGCGTCTGCATGACAGAGGTGGTGTCCAGTCCGCGTCCGTGACTCTCCGGAAACAC 984
 QY 308 Glu-----AlaValCysSerSer 313
 DB 985 GACATGACAGACTTCTTAGAAAAACCCACATGGGGACAGAGCCCTCTAAAGTTTTCAGCTCT 1044
 QY 314 GluProGluAsnValGlnAlaAspProGluAsnTyrThrSerIleLeuValThrTrpGlu 333
 DB 1045 CCACCATCCATGAAGGTGAGGCTCTGNAACGACGCGCACTGCAGGTGTCTCTGGAGC 1104
 QY 334 ArgProArgValValTyrAspThrMetIleGluLysPheAlaValLeuTyrGlnGlnLeu 353
 DB 1105 CAGCCGGAGACTATCTACACCCACCATCATGAAGTATCTCTTACAGCTGGACC 1164
 QY 354 AspGlyGluAspGlnThrLysHisGluPheLeuThrAspGlyTyrGlnAspLeu 371
 DB 1165 AAGAATGAGGACGAGAGAGAGAGAGCTTTTACAAAGGACAGGCGACAAAGACTTG 1218

RESULT 5

US-08-446-363-1
 ; Sequence 1, Application US/08446363
 ; Patent No. 5891700
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlessinger, Joseph
 ; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTHROSINE
 ; TITLE OF INVENTION: PHOSPHATASE-GAMMA
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PENNIE & EDMONDS
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/446,363
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/015,986
 ; FILING DATE: 10-FEB-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mistock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 7693-028
 ; TELEPHONE: 212-790-9090
 ; TELEFAX: 212-869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4338 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS

; LOCATION: 1..4335
 US-08-446-363-1

Alignment Scores: 3,736-69 Length: 4338
 Score: 648.50 Matches: 136
 Percent Similarity: 55.31% Conservative: 62
 Best Local Similarity: 37.99% Mismatches: 145
 Query Match: 32.75% Indels: 15
 DB: 2 Gaps: 2

US-09-983-000A-2 (1-374) x US-08-446-363-1 (1-4338)

QY 28 GlnGlnArgLysLeuValGluGluIleGlyTrpSerTyrThrGlyAlaLeuAsnGlnLys 47
 DB 148 CGCAGCGCAAGGCTTACGGCGACCCGCTACTGGGCTACTCTGGTGCTATGGTCTCTGAG 207
 QY 48 AsnTrpGlyLysLysTyrProThrCysAsnSerProLysGlnSerProIleAsnIleAsp 67
 DB 208 CACTGGGTCACTGTAGTGTCTAGTGTCTGGGAGCGCTCACCAGTCTCTTATTGACATTTA 267
 QY 68 GluAspLeuThrGlnValAsnValAsnLeuLysLysLysPheGlnGlyTrpAspLys 87
 DB 268 GACCAGTATGCGGTGTGGGGAAGAATACCAGGAACCTGCACTCGATGGCTTGCACAA 327
 QY 88 ThrSerLeuGluAsnThrPheIleHisAsnThrGlyLysThrValGluIleAsnLeuThr 107
 DB 328 GAGTCTTCTACAAACCTGGATGAAAAACACAGGGAACACAGTCCCATCTCTTGAA 387
 QY 108 AsnAspTyrArgValSerGlyLysValSerGluMetValPheLysAlaSerLysIleThr 127
 DB 388 GACGACTATTTTGTCTAGTGGAGTGTCTACTCTGGCAGATTTCMAAGCTGAGAAGTGGAA 447
 QY 128 PheHisTrpGlyLysCysAsnMetSerSerAspGlySerGluHisSerLeuGluGln 147
 DB 448 TTTCACTGGGGCCACAGCAAT---GGCTCAGCGGCTCTGAACACAGCATCAATGGCAGG 504
 QY 148 LysPheProLeuGluMetGlnIleTyrCysAspAlaAspArgPheSerSerPheGlu 167
 DB 505 AGGTTTCTGTGTAGATGAGATTTCTTTTACATCCAGATGACTTTTGACAGCTTTCA 564
 QY 168 GluAlaValLysGlyLysGlyLysLeuArgAlaLeuSerIleLeuPheGluValGlyThr 187
 DB 565 ACCGCAATTTCTGAGAACAGAAATAATCGAGGCCATGGCCATATTTTTCAGTCAATCGG 624
 QY 188 GluGluAsnLeuAspPheLysAlaIleAspGlyValGluSerValSerArgPheGly 207
 DB 625 AGGACCAATTTCTGCATCGATCTCTATTATCCACGGGTTGAAGGGTGTCTCATCATGAG 684
 QY 208 LysGlnAlaLeuAspProPheIleLeuLeuAsnLeuLeuProAsnSerThrAspLys 227
 DB 685 AAGAGACCTTTCTGGATCTCTTCTCTCGGGACCTCTCTGCTGCAATCTCTCTGGGCGAGC 744
 QY 228 TyrTyrIleTyrAsnGlySerLeuThrSerProCysThrAspThrValAspTrpIle 247
 DB 745 TATTATCGGTACACAGGTTCTTGACACACACACCGTGTAGCGAAATAGTGGAGTGGATA 804
 QY 248 ValPheLysAspThrValSerIleSerGluSerGlnLeuAlaValPheCysGluValLeu 267
 DB 805 GTCTTCGGGAGACCGCTCCCATCTCTTATCCATCAGCTTGAGGCTTTTATTTCATCTTC 864
 QY 268 ThrMetGlnGlnSerGlyTyrValMetLeuMetAspTyrIleGlnAsnAsnPheArgGlu 287
 DB 865 ACCACGGACGACGACGACCATGTCAAGTCGGTGAGTATCTGAGAAATAACTTTTCGACCA 924
 QY 288 GlnGlnTyrLysPheSerArgGlnValPheSerSerTyrThrGlyLysGluGluIleHis 307
 DB 925 CAGCAGCGTCTGCATGACAGAGGTGGTGTCCAAGTCCGCGCTCGTACTCTCTCGAACCAC 984
 QY 308 Glu-----AlaValCysSerSer 313
 DB 985 GACATGACAGACTTCTTAGAAAAACCCACTGGGGACAGAGCCCTCTAAAGTTTTCAGCTCT 1044

```
Qy 314 GluProGluAenValGlnAlaAspProGluAenTyrThrSerLeuValThrTyrGlu 333
Db 1045 CCACCCTCCATGAAGTGCAGCTCTGAACCAAGAGCGGACATGCTCTGAGC 1104
Qy 334 ArgProArgValValTyrAspThrMetIleGluLysPheAlaValLeuTyrGlnLeu 353
Db 1105 CAGCGGAGATATCTACCCACCACCATCATGAACATACATGATCTCTACAGCTGGACC 1164
Qy 354 AspGlyGluAenGlnThrLysHisGluPheLeuThrAspGlyTyrGlnAspLeu 371
Db 1165 AAGAATGAGGACGAGAGGAGAGCGTTTACAAAGGACGAGCAGCAAGACTTG 1218

RESULT 6
US-08-276-919-9
; Sequence 9, Application US/08276919
; Patent No. 5589579
; GENERAL INFORMATION:
; APPLICANT: Torczynski, Richard M.
; APPLICANT: Bollon, Arthur P.
; TITLE OF INVENTION: Lung Cancer Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RICHARDS, MEDLOCK & ANDREWS
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: TX
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276.919
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Neilligan, Mark C.
; REGISTRATION NUMBER: 36,389
; REFERENCE/DOCKET NUMBER: B35792
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 813 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..813
US-08-276-919-9

Alignment Scores:
Pred. No.: 4.7e-41 Length: 813
Score: 412.00 Matches: 93
Percent Similarity: 54.45% Conservative: 60
Best Local Similarity: 33.10% Mismatches: 108
Query Match: 20.81% Indels: 20
DB: 1 Gaps: 6

US-09-983-000A-2 (1-374) x US-08-276-919-9 (1-813)
Qy 38 TrpSerTyrThrGlyAlaLeuAenGlnLysAenTyrGlyLysValTyrProThrCysAen 57
Db 7 TGGACTTAATTGGTCTGCTGGGAGAAATAGTGGTCCAAAGAGTACCGCTGTGTGGG 66
Qy 58 SerProLysGlnSerProIleAenIleAspGluAenLeuThrGlnValAenValAenLeu 77
Db 67 GGCCTGTGCTGCCAGTCCCGCCATAGACCTGCACAGTGCATCCTCCAGATGATGACGCCAGCCTC 126
```

```
Qy 78 LysLysLeuLysPheGlnGlyTyrAspLysThrSerLeuGluAenThrPheIleHisAen 97
Db 127 ACGCCCTCAGTTCAGAGTCAATCTGTCTGCCAAGCAAGAGTTTCTCTCACCACCA 186
Qy 98 ThrGlyLysThrValGluIleAenLeuThrAenAspTyrArgValSerGlyGlyValSer 117
Db 187 ANTGGCCATTGAGTGAAGCTGACCTGCCCTCGGACATGCACATCCAGGCGCTCAGTCT 246
Qy 118 GluMetValPheLysAlaSerLysIleThrPheHisTyrGlyLysCysAenMetSerSer 137
Db 247 CGC-----TACAGTGCACGACGCTGCACCTGCCTGCTGGGGGAACCCGAAT---GACCCG 297
Qy 138 AspGlySerGluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCys 157
Db 298 CACGGCTCTGAGCACACCGTCAGCGGACAGCATTTCGCCCGCGGAGTGCACATTGTCCAT 357
Qy 158 PheAspAlaAspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArg 177
Db 358 TATAACTCAGACCTTTATCTCTGACGCCAGCAGCTGCCAGCAACAAGTCAGAGGCTCGCT 417
Qy 178 AlaLeuSerIleLeuPheGluValGlyThrGluGluAenLeuAenPheLysAlaIleIle 197
Db 418 GTCCTGGCTGTCTCTCATTTGATGGGCTCC---TTCAATCGTCTCTATGACAAGATCTTC 474
Qy 198 AspGlyValGluSerValSerArgPheGlyLysGlnAlaLeuAenAspProPheIleLeu 217
Db 475 AGTCACCTTCAACATGTAAGTACAAAGCGGACGAGCATTCGTCGCCGGATTCAACATT 534
Qy 218 LeuAenLeuLeuProAenSerThrAspLysTyrTyrIleTyrAenGlySerLeuThrSer 237
Db 535 GAAGAGCTGCTTCGGAGAGGACCGCTGATATTTACCGCTACCGGGGGTCCCTGACCACCA 594
Qy 238 ProProCysThrAspThrValAspTyrIleValPheLysAspThrValSerIleSerGlu 257
Db 595 CCCCTTGCACACCCCACTGTCTCTGCACAGTTTTTCGAAACCCCGTGCATAATTTCACG 654
Qy 258 SerGlnLeuAlaValPheCysGluValLeuThrMetGlnSerGlyTyrValMetLeu 277
Db 655 GAGCAGCTG-----CTGGCTTTGGAGACAGCCCTGTACTGCACACAC 696
Qy 278 MetAsp-----TyrLeuGlnAenAenPheArgGluGlnGlnTyrLysPhe 292
Db 697 ATGACGACCCCTTCCCCCAGAGAAATGATCAACAACTTCCGGCAGGTCAGAGTTCCAT 756
Qy 293 SerArgGlnValPheSerSerTyrThrGlyLysGluGluIleHisGluAlaValCysSer 312
Db 757 GAGAGGCTGGTATACACCTCTCTCTCCCAAGTGCAA-----GTCTGTACT 801
Qy 313 Ser 313
Db 802 GCG 804

RESULT 7
US-08-776-088-12
; Sequence 12, Application US/08776088
; Patent No. 5773579
; GENERAL INFORMATION:
; APPLICANT: Torczynski, Richard M.
; APPLICANT: Bollon, Arthur P.
; TITLE OF INVENTION: Lung Cancer Marker
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SIDLEY & AUSTIN
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: TX
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```


; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/776,088
 ; FILING DATE: 19-JUL-95
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Eugenia S. Hansen
 ; REGISTRATION NUMBER: 31,966
 ; REFERENCE/DOCKET NUMBER: 10365/05011
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 214-981-3300
 ; TELEFAX: 214-981-3400
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 813 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..813
 ; US-08-776-088-12

Alignment Scores:

Pred. No.:	4.7e-41	Length:	813
Score:	412.00	Matches:	93
Percent Similarity:	54.45%	Conservative:	60
Best Local Similarity:	33.10%	Mismatches:	108
Query Match:	20.81%	Indels:	20
DB:	1	Gaps:	6

US-09-983-000A-2 (1-374) x US-08-776-088-12 (1-813)

QY 38 TrpSerThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysValProThrCysAsn 57
 DB 7 TGGACTATTTGGTCTCATGGGAGGATAGTGTCCAGAGTACCGGCGGTGGG 66
 QY 58 SerProLysGlnSerProLeuAsnLeuThrGlnValAsnValAsnLeu 77
 DB 67 GCGCTGCTGCATCCCCATAGACGTGCACATCTCTCCATGATGACGCCAGCTC 126
 QY 78 LysLysLeuLysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleHisAsn 97
 DB 127 ACGCCCTCGAGTCCAGGCTACAACTGTCTGCGCAACAGCAGTTTCTCTGACCAAC 186
 QY 98 ThrGlyLysThrValGluLeuAsnLeuThrAsnAspThrArgValSerGlyValSer 117
 DB 187 AATGGCCATTTCAGTGAAGCTGAACCTGCGGACATGCACATCCAGGGCCCTCCAGTCT 246
 QY 118 GluMetValPheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSer 137
 DB 247 CGG-----TACAGTGCACGAGCTGCACCTGCTGGGGAACCCGAT---GACCCG 297
 QY 138 AspGlySerGluHisSerLeuGluGlnLysPheProLeuGluMetGlnIleThrCys 157
 DB 298 CACGGCTCTGAGCACACCGCTCAGCGGACAGCACTTCGCGCGGAGCTGCACATTGTCCAT 357
 QY 158 PheAspAlaAspArgPheSerSerPheGluGluAlaValLysGlyLysLysLeuArg 177
 DB 358 TATAACTCAGACTTTATCTGACGCCAGCAGCTGCCAGCAACCAAGTCAGAGGCCCTCGCT 417
 QY 178 AlaLeuSerIleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIle 197
 DB 418 GTCTGGCTGTTCTATTGAGATGGGCTCC---TTCAATCCGCTCTATGACAGATCTTC 474
 QY 198 AspGlyValGluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeu 217
 DB 475 AGTCACCTTCAACATGTAAAGTACAAAGCCAGGAGCATTCGTCGCCGGATTCAACATT 534
 QY 218 LeuAsnLeuLeuProAsnSerThrAspLysThrIleThrAsnGlySerLeuThrSer 237
 DB 535 GAAGAGCTGCTTCCGAGAGGAGCCGCTGAATATTACCGCTACCGGGGGTCCCTGACACA 594

QY 238 ProProCysThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGlu 257
 DB 595 CCCCTTGGCAACCCCACTGTCTGGACAGTTTTCCGAACCCCGTGCATAATTTCCAG 654
 QY 258 SerGlnLeuAlaValPheCysGluValLeuThrMetGlnGlnSerGlyThrValMetLeu 277
 DB 655 GAGCAGCTG-----CTGGCTTTGGAGACAGCCCTGTACTGTGCACACAC 696
 QY 278 MetAsp-----TyrLeuGlnAsnAsnPheArgGluGlnGlnGlnIleThrLysPhe 292
 DB 697 ATGGACGACCTTCCCGCAGAGAATGATCAACACTTCCGCGAGGTCCAGAGTTCGAT 756
 QY 293 SerArgGlnValPheSerSerThrGlyLysGluGluIleHisGluAlaValCysSer 312
 DB 757 GAGAGGCTGGTATACACCTCTCTCCCAAGTGCAA-----GTCTGTACT 801
 QY 313 Ser 313
 DB 802 GCG 804

RESULT 8
 PCT-US95-09145A-12
 ; Sequence 12, Application PC/TUS9509145A
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: Lung Cancer Marker
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: RICHARDS, MEDLOCK & ANDREWS
 ; STREET: 1201 Elm Street, Suite 4500
 ; CITY: Dallas
 ; STATE: TX
 ; COUNTRY: US
 ; ZIP: 75270-2197
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/09145A
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: John A. Harre
 ; REGISTRATION NUMBER: 37,345
 ; REFERENCE/DOCKET NUMBER: B35792CIPPCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 214-939-4500
 ; TELEFAX: 214-939-4600
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 813 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..813
 ; PCT-US95-09145A-12

Alignment Scores:	4.7e-41	Length:	813
Pred. No.:	412.00	Matches:	93
Score:	54.45%	Conservative:	60
Percent Similarity:	33.10%	Mismatches:	108
Best Local Similarity:	20.81%	Indels:	20
Query Match:	5	Gaps:	6
DB:			

US-09-983-000A-2 (1-374) x PCT-US95-09145A-12 (1-813)

```
QY 38 TrpSerTyrThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsn 57
Db 7 TGGACTTATTTGGTCTGTGAGGGGAGAAATAGCTGTCACAGAAAGTACCCGTCGTGTGGG 66
QY 58 SerProLysGlnSerProIleAsnIleAspGluLeuAspLeuThrGlnValAsnValAsnLeu 77
Db 67 GGCCTGCTGCAGTCCCTCCATAGACTGCACAGTGCATCTCCAGTATGACGCCAGCTC 126
QY 78 LysLysLeuLysPheGlnGlyTyrPaspLysThrSerLeuGluAsnThrPheIleHis 97
Db 127 ACGCCCTCGAGTTCACAGGCTACAATCTGTGCCAACAGCAGTTCCTCTGACCAAC 186
QY 98 ThrGlyLysThrValGluLeuAsnLeuThrAsnAspTyrArgValSerGlyGlyValSer 117
Db 187 AATGGCCATTCAGTGAAGTGAACCTGCCCTCGACATGCACATCCAGGCTCCAGTCT 246
QY 118 GluMetValPheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSer 137
Db 247 CGC-----TACAGTGCACGACCTGCATCTGCTGGGGGAACCCGAAT---GACCCG 297
QY 138 AspGlySerGluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCys 157
Db 298 CACGGCTCTGAGCACACCTTCGCGGAGCAGCAGTTCGCGCGGAGCTGCACATTTGTCAT 357
QY 158 PheAspAlaAspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLeuArg 177
Db 358 TATAACTCAGACCTTTATCTCGACGCCAGCAGTCTGCCAGCAACAAGTCAGAGGCTCGCT 417
QY 178 AlaLeuSerIleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIle 197
Db 418 GTCCCTGGCTGTCTCTCATTTGAGATGGCTCC---TTCAATCGTCTCTATGACAGATCTTC 474
QY 198 AspGlyValGluSerValSerArgPheGlyLysGlnAlaAlaLeuAspPhePheIleLeu 217
Db 475 AGTCACTTCAACATTAAGTAAAGTACAAAGCCAGGAGCATTCGTCGCCGGATTCACATT 534
QY 218 LeuAsnLeuLeuProAsnSerThrAspLysTyrIleTyrAsnGlySerLeuThrSer 237
Db 535 GAAAGCTGCTTCCGGAGGAGGACCGCTGAATATTACCGCTACCGGGGGTCCCTGACCA 594
QY 238 ProProCysThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGlu 257
Db 595 CCCCTCTGCAACCCACATGTGCTCTGGACAGTTTTCGAAACCCCGTGCATAATTTCCAG 654
QY 258 SerGlnLeuAlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeu 277
Db 655 GAGCAGCTG-----CTGGCTTTGGAGACAGCCCTGTACTGACACAC 696
QY 278 MetAsp-----TyrLeuGlnAsnAsnAspPheArgGluGlnGlnTyrIlePhe 292
Db 697 ATGACGACCCCTTCCCGAGAGAAATGATCAACAATTCGCGGAGTCCAGAAATTCGAT 756
QY 293 SerArgGlnValPheSerSerTyrThrGlyLysGluGluIleHisGluAlaValCysSer 312
Db 757 GAGAGCTGTATACACCTCTCTCTCCCAAGTGCAA-----GTCTGTACT 801
QY 313 Ser 313
Db 802 GCG 804
RESULT 9
US-08-776-088-14
; Sequence 14, Application US/08776088
; Patent No. 5773579
; GENERAL INFORMATION:
; APPLICANT: Torczynski, Richard M.
; APPLICANT: Bollon, Arthur P.
; TITLE OF INVENTION: Lung Cancer Marker
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SIDLEY & AUSTIN
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
```

```
STATE: TX
COUNTRY: US
ZIP: 75270-2197
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,088
FILING DATE: 19-JUL-95
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Eugenia S. Hansen
REGISTRATION NUMBER: 31,966
REFERENCE/DOCKET NUMBER: 10365/05011
TELECOMMUNICATION INFORMATION:
TELEPHONE: 214-981-3300
TELEFAX: 214-981-3400
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 822 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..822
US-08-776-088-14
Alignment Scores:
Pred. No.: 4,78e-41 Length: 822
Score: 412.00 Matches: 93
Percent Similarity: 54.45% Conservative: 60
Best Local Similarity: 33.10% Mismatches: 108
Query Match: 20.81% Indels: 20
DB: 1 Gaps: 6
US-09-983-000A-2 (1-374) x US-08-776-088-14 (1-822)
QY 38 TrpSerTyrThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsn 57
Db 7 TGGACTTATTTGGTCTGTGAGGGGAGAAATAGCTGTCACAGAAAGTACCCGTCGTGTGGG 66
QY 58 SerProLysGlnSerProIleAsnIleAspGluLeuAspLeuThrGlnValAsnValAsnLeu 77
Db 67 GGCCTGCTGCAGTCCCTCCATAGACTGCACAGTGCATCTCTCCAGTATGACGCCAGCTC 126
QY 78 LysLysLeuLysPheGlnGlyTyrPaspLysThrSerLeuGluAsnThrPheIleHis 97
Db 127 ACGCCCTCGAGTTCACAGGCTACAATCTGTCTGCCAACAGCAGTTCCTCTGACCAAC 186
QY 98 ThrGlyLysThrValGluLeuAsnLeuThrAsnAspTyrArgValSerGlyGlyValSer 117
Db 187 AATGGCCATTCAGTGAAGTGAACCTGCCCTCGACATGCACATCCAGGCTCCAGTCT 246
QY 118 GluMetValPheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSer 137
Db 247 CGC-----TACAGTGCACGACCTGCATCTGCTGGGGGAACCCGAAT---GACCCG 297
QY 138 AspGlySerGluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCys 157
Db 298 CACGGCTCTGAGCACACCTTCGCGGAGCAGCAGTTCGCGCGGAGCTGCACATTTGTCAT 357
QY 158 PheAspAlaAspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLeuArg 177
Db 358 TATAACTCAGACCTTTATCTCGACGCCAGCAGTCTGCCAGCAACAAGTCAGAGGCTCGCT 417
QY 178 AlaLeuSerIleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIle 197
Db 418 GTCCCTGGCTGTCTCTCATTTGAGATGGCTCC---TTCAATCGTCTCTATGACAGATCTTC 474
```



```

; APPLICANT: Bollon, Arthur P.
; TITLE OF INVENTION: Lung Cancer Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RICHARDS, MEDLOCK & ANDREWS
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: TX
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,919
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelligan, Mark C.
; REGISTRATION NUMBER: 36,389
; REFERENCE/DOCKET NUMBER: B35792
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 986 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..975
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 895..906
; OTHER INFORMATION: /note= "phosphorylation site
; OTHER INFORMATION: recognized by protein C kinase and other kinases"
US-08-276-919-3

Alignment Scores:
Pred. No.: 6,418-41 Length: 986
Score: 412.00 Matches: 93
Percent Similarity: 54.45% Conservative: 60
Best Local Similarity: 33.10% Mismatches: 108
Query Match: 20.81% Indels: 20
DB: 1 Gaps: 6

US-09-983-000a-2 (1-374) x US-08-276-919-3 (1-986)

Qy 38 TrpSerThrGluValAlaLeuAsnGlnLysAsnTrpGlyLysValSerProThrCysAsn 57
Db 7 TGGACTTATTTGGTCTGATGGGAGATAGTGGTCCAGAGATACCGTGTGGG 66

Qy 58 SerProLysGlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeu 77
Db 67 GGCCGTGCTGACGTCCCGCCATAGACCTGCACAGTGCATCTCCAGTATGACGCCAGCCTC 126

Qy 78 LysLysLeuLysPheGlnGlyTrpAspLysThrSerLeuGluAsnThrPheIleHisAsn 97
Db 127 ACGCCCTCGAGTCCAAAGGCTACATCTGTCTGCCAACAAGCAGTTTCTCTGACCAAC 186

Qy 98 ThrGlyLysValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSer 117
Db 187 ATGGCCATTCAGTAAGCTGACTGCTCCGACATGCATCCAGGCTCCAGTCT 246

Qy 118 GluMetValPheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSer 137
Db 247 CGC-----TACAGTGCCACGACGACGTGCACCTGCGACTGGGGAAACCCGAAT---GACCCG 297

```

```

Qy 138 AspGlySerGluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCys 157
Db 298 CACGGCTCTGAGCACACCGTCAGCGGACAGCACTTCGCCGCCGAGCTGCACATTGTCCAT 357

Qy 158 PheAspAlaAspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArg 177
Db 358 TATAACTCAGACCTTTATCTCGAGCCGACGACTGCCAGCAACAAGTCAGAAGGCTCGCT 417

Qy 178 AlaLeuSerIleLeuPheGluValGlyThrGluAsnLeuAspPheLysAlaIleLeu 197
Db 418 GTCTGTGCTGTCTCATTTGAGATGGGCTCC---TTCAATCCGTCCTATGACAAGATCTTC 474

Qy 198 AspGlyValGluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeu 217
Db 475 AGTCACCTTCAACATGTAAAGTCAAAAGGCCAGAGACATTCGCCGGGATTCACATT 534

Qy 218 LeuAsnLeuLeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSer 237
Db 535 GAAGAGCTGCTTCGGAGAGAGCCGCTGAATATTACCGCTACCGGGGTCCTCGACCACA 594

Qy 238 ProProCysThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGlu 257
Db 595 CCCCTTGCACACCCCACTGTCTCTGCACAGTTTTTCGAAACCCCGTGCATAATTCCAG 654

Qy 258 SerGlnLeuAlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeu 277
Db 655 GAGCAGCTG-----CTGGCTTTGGAGACAGCCCTGTACTGCACACAC 696

Qy 278 MetAsp-----TyrLeuGlnAsnAsnPheArgGluGlnGlnTyrLysPhe 292
Db 697 ATGACGACCCCTTCCCCCAGAGAAATGATCAACAACCTTCGGCAGGTCAGAGTTCCGAT 756

Qy 293 SerArgGlnValPheSerSerTyrThrGlyLysGluGluLeuHisGluAlaValCysSer 312
Db 757 GAGAGGCTGGTATACACCTCTCTCCCAAGTGCAA-----GTCTGTACT 801

Qy 313 Ser 313
Db 802 GCG 804

RESULT 12
US-08-276-919-12
; Sequence 12, Application US/08276919
; Patent No. 5589579
; GENERAL INFORMATION:
; APPLICANT: Torczynski, Richard M.
; APPLICANT: Bollon, Arthur P.
; TITLE OF INVENTION: Lung Cancer Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RICHARDS, MEDLOCK & ANDREWS
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: TX
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,919
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelligan, Mark C.
; REGISTRATION NUMBER: 36,389
; REFERENCE/DOCKET NUMBER: B35792
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600
; INFORMATION FOR SEQ ID NO: 12:

```

SEQUENCE CHARACTERISTICS:
 LENGTH: 986 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..975
 US-08-276-919-12

Alignment Scores:
 Pred. No.: 6,41e-41 Length: 986
 Score: 412.00 Matches: 93
 Percent Similarity: 54.45% Conservative: 60
 Best Local Similarity: 33.10% Mismatches: 108
 Query Match: 20.81% Indels: 20
 DB: 1 Gaps: 6

US-09-983-000A-2 (1-374) x US-08-276-919-12 (1-986)

```

QY 38 TTPSerTyrThrGlyAlaLeuAenGlnLysAenTTPGlyLysTyrProThrCysAen 57
DB 7 TGGACTTATTGGTCTGATGGGAGATAGCTGTCCAGAGTACCCGTCGTGGG 66

QY 58 SerProLysGlnSerProIleAenLysLeuThrGlnValAenValAenLeu 77
DB 67 GGCCTGCTGACGTCCCTCCATAGCTGCACATGATCTCCATGATGACGCCAGCTC 126

QY 78 LysLysLeuLysPheGlnGlyTrpAspLysThrSerLeuGlnAenThrPheIleHisAen 97
DB 127 ACGCCCTCGAGTTCAGAGCTTACATCTGTGTCACACAGCAGTTCCTCTGACCAAC 186

QY 98 ThrGlyLysThrValGluIleAenLeuThrAenAspTyrArgValSerGlyValSer 117
DB 187 AATGCCATCTAGTGAAGCTGAACCTGCTCCGACATGCATCCAGGGCTCCAGTCT 246

QY 118 GluMetValPheLysAlaSerLysIleThrPheHisTTPGlyLysCysAenMetSerSer 137
DB 247 CGC-----TACAGTGCCACGCAGCTGCACCTGCCTGGGGAACCCGAT---GACCCG 297

QY 138 AspGlySerGluHisSerLeuGluGlnLysPheProLeuGluMetGlnIleTyrCys 157
DB 298 CACGGCTCTGAGCACACCGCTCAGCGCAGCAGCACTTCGCGCGAGCTGCACATTCCTCCAT 357

QY 158 PheAspAlaAspArgPheSerPheGluGluAlaValLysGlyLysLeuArg 177
DB 358 TATAACTCAGACCTTATCTCAGCCGACGACTGCCAGCAACAGTCAGAGGCTCGCT 417

QY 178 AlaLeuSerIleLeuPheGluValGlyThrGluGlnAenLeuAspPheLysAlaIleIle 197
DB 418 GTCTGTGGCTGTCTCATTTGAGATGGCTCC---TTCAATCCGCTCTATGACAGATCTTC 474

QY 198 AspGlyValGluSerValSerArgPheGlyLysGlnAlaLeuAspPropheIleLeu 217
DB 475 AGTCACCTTCAACATGTAAAGTACAAAGGCCAGGAAGCATTCGTCGCGGATTCACCAT 534

QY 218 LeuAenLeuLeuProAenSerThrAspLysTyrIleTyrAenGlySerLeuSer 237
DB 535 GAAGAGCTGTCTCCGAGAGGACCCCTGAATATATACCGCTACCGGGGGTCCCTGACCA 594

QY 238 ProProCysThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGlu 257
DB 595 CCCCCTTGCAACCCCACTGTGCTCTGGACAGTTTTCGGAACCCCGTGCMAATTTCCAC 654

QY 258 SerGlnLeuAlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeu 277
DB 655 GAGCAGCTG-----CTGGCTTTGGAGACAGCCCTCTACTGCACACAC 696

QY 278 MetAsp-----TyrLeuGlnAenAenPheArgGluGlnGlnTyrLysPhe 292
DB 697 ATGGACGACCCCTTCCCCCAGAGAAATGATCAACACTTCCGCGCAGGTCCTCCAGAGTTCCAT 756

```

```

QY 293 SerArgGlnValPheSerSerTyrThrGlyLysGluGluIleHisGluAlaValCysSer 312
DB 757 GAGAGCGTGTATACACCTCTCTCCCAAGTGCAA-----GTCTGTACT 801

QY 313 Ser 313
DB 802 GCG 804

```

RESULT 13

US-08-776-088-3
 ; Sequence 3, Application US/08776088
 ; Patent No. 5773579
 ; GENERAL INFORMATION:
 ; APPLICANT: Torczynski, Richard M.
 ; APPLICANT: Bollon, Arthur P.
 ; TITLE OF INVENTION: Lung Cancer Marker
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SIDLEY & AUSTIN
 ; CITY: Dallas
 ; STATE: TX
 ; COUNTRY: US
 ; ZIP: 75270-2197
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/776.088
 ; FILING DATE: 19-JUL-95
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Eugenia S. Hansen
 ; REGISTRATION NUMBER: 31,966
 ; REFERENCE/DOCKET NUMBER: 10365/05011
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 214-981-3300
 ; TELEFAX: 214-981-3400
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 986 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..975
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 895..906
 ; OTHER INFORMATION: /note= "phosphorylation site
 ; OTHER INFORMATION: recognized by protein C kinase and other kina..."
 US-08-776-088-3

Alignment Scores:
 Pred. No.: 6,41e-41 Length: 986
 Score: 412.00 Matches: 93
 Percent Similarity: 54.45% Conservative: 60
 Best Local Similarity: 33.10% Mismatches: 108
 Query Match: 20.81% Indels: 20
 DB: 1 Gaps: 6

US-09-983-000A-2 (1-374) x US-08-776-088-3 (1-986)

```

QY 38 TTPSerTyrThrGlyAlaLeuAenGlnLysAenTTPGlyLysTyrProThrCysAen 57
DB 7 TGGACTTATTGGTCTGATGGGAGATAGCTGTCCAGAGTACCCGTCGTGGG 66

QY 58 SerProLysGlnSerProIleAenLysLeuThrGlnValAenValAenLeu 77

```

Db 67 GGCCTGCTGCAGTCCCCCATAGACCTGCACAGTGCATCTCCAGTATGACCCAGCCTC 126
Qy 78 LysLysLeuLysPheGlnGlyThrPheLysThrSerLeuGluAsnThrPheIleHisAsn 97
Db 127 ACGTCCCTCGAGTTCGAAGGCTACAATCTGTCTGCAACAAGCAGATTTCTCTCGACCAAC 186
Qy 98 ThrGlyLysThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSer 117
Db 187 AATGGCCATTGATGAAGTGAACCTGCCCTGCATGCACATGCCAGGGCTCCAGTCT 246
Qy 118 GluMetValPheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSer 137
Db 247 CGC-----TACAGTGCACCGAGCTGCATCTGGGGAACCCGAAT---GACCCG 297
Qy 138 AspGlySerGluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCys 157
Db 298 CACGGCTCTGAGCACACCGTTCAGCGGACAGCAGCTTCGGCGCGAGCTGCACATTGTCCAT 357
Qy 158 PheAspAlaAspArgPheSerPheGluGluAlaValLysGlyLysGlyLysLeuArg 177
Db 358 TATAACTCAGACCTTTATCTGACGCCAGCAGCTGCCAGCAACAAGTCAAGAGGCTCGCT 417
Qy 178 AlaLeuSerIleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIle 197
Db 418 GTCCTGGCTGTCTCATTTGAGATGGGCTCC---TTCAATCCGCTCTATGACAAATCTTC 474
Qy 198 AspGlyValGluSerValSerArgPheGlyLysGlnAlaLeuAspPropheIleLeu 217
Db 475 AGTCACCTTCAACATGTAAGTACAAAGGCCAGGACGATTCGTCGCCGGATTCAACATT 534
Qy 218 LeuAsnLeuLeuProAsnSerThrAspLysTyrIleTyrAsnGlySerLeuThrSer 237
Db 535 GAAGAGCTGCTTCGGAGAGGACCGCTGAATATTACCGCTACCGGGGGTCCCTGACACACA 594
Qy 238 ProProCysThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGlu 257
Db 595 CCCCCTTGCAACCCACCTGTGCTCTGGACAGTTTTCGGAACCCCGTGCATTTCCAG 654
Qy 258 SerGlnLeuAlaValPheCysGluValLeuThrMetGlnSerGlyTyrValMetLeu 277
Db 655 GAGTAGCTG-----CTGGCTTTGGAGACAGCCCTGTACTGTGCACACAC 696
Qy 278 MetAsp-----TyrLeuGlnAsnAsnPheArgGluGlnGlnTyrLysPhe 292
Db 697 ATGACGACCTTCCCCAGAGAATGATCAACAACCTTCGGCGAGCTCCAGAAGTTCAT 756
Qy 293 SerArgGlnValPheSerSerTyrThrGlyLysGluGluIleHisGluAlaValCysSer 312
Db 757 GAGAGGCTGGTATACACCTCTCTCCCAAGTGCAA-----GTCTGTACT 801
Qy 313 Ser 313
Db 802 GCG 804

RESULT 14

US-08-776-088-17
; Sequence 17, Application US/08776088
; Patent No. 5773579

GENERAL INFORMATION:

; APPLICANT: Torczynski, Richard M.
; APPLICANT: Bollon, Arthur P.
; TITLE OF INVENTION: Lung Cancer Marker
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: SIDLEY & AUSTIN
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: TX
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,088
FILING DATE: 19-JUL-95
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Eugenia S. Hansen
REGISTRATION NUMBER: 31,966
REFERENCE/DOCKET NUMBER: 10365/05011
TELECOMMUNICATION INFORMATION:
TELEPHONE: 214-981-3300
TELEFAX: 214-981-3400
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 986 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..975
US-08-776-088-17
Alignment Scores:
Pred. No.: 6,41e-41 Length: 986
Score: 412.00 Matches: 93
Percent Similarity: 54.45% Conservative: 60
Best Local Similarity: 33.10% Mismatches: 108
Query Match: 20.81% Indels: 20
DB: 1 Gaps: 6
US-09-983-000A-2 (1-374) x US-08-776-088-17 (1-986)
Qy 38 TrpSerTyrThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsn 57
Db 7 TGGACTTATTTTGGTCTGTGATGGGAGATAGTGGTCCAGGAAGTACCCTGCTGTGGG 66
Qy 58 SerProLysGlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeu 77
Db 67 GGCCTGCTGCAGTCCCATAGACCTGCACAGTGCATCTCCAGTATGACCCAGCCTC 126
Qy 78 LysLysLeuLysPheGlnGlyThrAspLysThrSerLeuGluAsnThrPheIleHisAsn 97
Db 127 ACGCCCTCGAGTTCGAAGCTACAATCTGTCTGCCAACAAGCAGTTCCTCTGACCAAC 186
Qy 98 ThrGlyLysThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyGlyValSer 117
Db 187 AATGGCCATTGATGAAGTGAACCTGCCCTCGGACATGCACATCCAGGGCTCCAGTCT 246
Qy 118 GluMetValPheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSer 137
Db 247 CGC-----TACAGTGCACCGAGCTGCATCTGGGGAACCCGAAT---GACCCG 297
Qy 138 AspGlySerGluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCys 157
Db 298 CACGGCTCTGAGCACACCGTTCAGCGGACAGCAGCTTCGGCGCGAGCTGCACATTGTCCAT 357
Qy 158 PheAspAlaAspArgPheSerPheGluGluAlaValLysGlyLysGlyLysLeuArg 177
Db 358 TATAACTCAGACCTTTATCTGACGCCAGCAGCTGCCAGCAACAAGTCAAGAGGCTCGCT 417
Qy 178 AlaLeuSerIleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIle 197
Db 418 GTCCTGGCTGTCTCATTTGAGATGGGCTCC---TTCAATCCGCTCTATGACAAATCTTC 474
Qy 198 AspGlyValGluSerValSerArgPheGlyLysGlnAlaLeuAspPropheIleLeu 217
Db 475 AGTCACCTTCAACATGTAAGTACAAAGGCCAGGACGATTCGTCGCCGGATTCAACATT 534
Qy 218 LeuAsnLeuLeuProAsnSerThrAspLysTyrIleTyrAsnGlySerLeuThrSer 237

Db 535 GAAGAGCTGCTCCCGGAGAGGACCGCTGAATATTACCGCTACCGGGGGTCCCTGACACCA 594
Qy 238 ProProCysThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGlu 257
Db 595 CCCCCTTGCAACCCCACTGCTCTGGACAGTTTCCGAAACCCCGTGCMAATTTCCACAG 654
Qy 258 SerGlnLeuAlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeu 277
Db 655 GAGCAGCTG-----CTGGCTTTGGAGACAGCCCTGTACTGCACACAC 696
Qy 278 MetAsp-----TyrLeuGlnAsnAsnPheArgGluGlnGlnTyrIlysPhe 292
Db 697 ATGGACGACCCCTCCCGCAGAGAAATGATCAACAACCTTCCGGCAGTCCAGAAAGTTCCAT 756
Qy 293 SerArgGlnValPheSerTyrThrGlyLysGluGluIleHisGluAlaValCysSer 312
Db 757 GAGAGGCTGGTATACACCTCTCTCCCAAGTGCAA-----GTCTGTACT 801
Qy 313 Ser 313
Db 802 GCG 804

RESULT 15

PCT-US95-09145A-3
; Sequence 3, Application PC/TUS9509145A
; GENERAL INFORMATION:
; APPLICANT: Lung Cancer Marker
; TITLE OF INVENTION: 22
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RICHARDS, MEDLOCK & ANDREWS
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: TX
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09145A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: John A. Harre
; REGISTRATION NUMBER: 37,345
; REFERENCE/DOCKET NUMBER: B35792CIPPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 986 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..975
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 895..906
; OTHER INFORMATION: /note= "phosphorylation site
; OTHER INFORMATION: recognized by protein C kinase and other kina..."
PCT-US95-09145A-3

Alignment Scores:
Pred. No.: 6.41e-41 Length: 986
Score: 412.00 Matches: 93
Percent Similarity: 54.45% Conservative: 60

Best Local Similarity: 33.10% Mismatches: 108
Query Match: 20.81% Indels: 20
DB: 5 Gaps: 6
US-09-983-000A-2 (1-374) x PCT-US95-09145A-3 (1-986)
Qy 38 TrpSerTyrThrGlyAlaLeuAsnGlnLysAsnTrpGlyLysLysTyrProThrCysAsn 57
Db 7 TGGACTTATTTGGTCTGCTGAGGAGAAATAGCTGGTCCAAAGAGTACCCGCTGCTGGG 66
Qy 58 SerProLysGlnSerProIleAsnIleAspGluAspLeuThrGlnValAsnValAsnLeu 77
Db 67 GSCCTGTGTCAGTCCCCCATAGACGTGCACAGTGCACATCTCCAGATATGATGAGCCGCTC 126
Qy 78 LysLysLeuLysPheGlnGlyTyrAspLysThrSerLeuGluAsnThrPheIleHisAsn 97
Db 127 AGCCCCCTCGAGTGTCCAGGCTCAATCTGTCTCCCAACAAGCAGTTTCTCTGACCAAC 186
Qy 98 ThrGlyLysThrValGluIleAsnLeuThrAsnAspTyrArgValSerGlyLysValSer 117
Db 187 AATGCCCATTCAGTGAAGCTGAACCTGCCCTCGACATGCACATCCAGGGGCTCCAGTCT 246
Qy 118 GluMetValPheLysAlaSerLysIleThrPheHisTrpGlyLysCysAsnMetSerSer 137
Db 247 CGC-----TACAGTGCACGACGCTGCACCTGCACCTGGGGGAAACCCGAAT---GACCCG 297
Qy 138 AspGlySerGluHisSerLeuGluGlyGlnLysPheProLeuGluMetGlnIleTyrCys 157
Db 298 CACGGCTCTGAGCACACCGCTCAGCGGACAGCAGCTTCCGCCGCGAGCTGCACATGTCTCAT 357
Qy 158 PheAspAlaAspArgPheSerSerPheGluGluAlaValLysGlyLysGlyLysLeuArg 177
Db 358 TATAACTCAGACCTTTATCTCTGACGCGCAGCACTGCCAGCAACAGTCAGAGGCTCGCT 417
Qy 178 AlaLeuSerIleLeuPheGluValGlyThrGluGluAsnLeuAspPheLysAlaIleIle 197
Db 418 GTCCCTGGCTGTCTCTCATTTGAGATGGGCTCC---TTCAATCCGCTCTATGACACAGATCTTC 474
Qy 198 AspGlyValGluSerValSerArgPheGlyLysGlnAlaLeuAspProPheIleLeu 217
Db 475 AGTCACCTTCAACATGTAAAGTCAAAAGGCCAGGAGCATTCGTCGCCGGGATTCACAT 534
Qy 218 LeuAsnLeuLeuProAsnSerThrAspLysTyrTyrIleTyrAsnGlySerLeuThrSer 237
Db 535 GAAGAGCTGCTTCGGAGAGACCGCTGAATATTACCGCTACCGGGGGTCCCTGACACCA 594
Qy 238 ProProCysThrAspThrValAspTrpIleValPheLysAspThrValSerIleSerGlu 257
Db 595 CCCCCTTGCAACCCCACTGCTCTGGACAGTTTTCGAAACCCCGTGCMAATTTCCACAG 654
Qy 258 SerGlnLeuAlaValPheCysGluValLeuThrMetGlnGlnSerGlyTyrValMetLeu 277
Db 655 GAGCAGCTG-----CTGGCTTTGGAGACAGCCCTGTACTGCACACAC 696
Qy 278 MetAsp-----TyrLeuGlnAsnAsnPheArgGluGlnGlnTyrIlysPhe 292
Db 697 ATGGACGACCCCTCCCGCAGAGAAATGATCAACAACCTTCCGGCAGTCCAGAAAGTTCCAT 756
Qy 293 SerArgGlnValPheSerTyrThrGlyLysGluGluIleHisGluAlaValCysSer 312
Db 757 GAGAGGCTGGTATACACCTCTCTCCCAAGTGCAA-----GTCTGTACT 801
Qy 313 Ser 313
Db 802 GCG 804

Search completed: August 30, 2005, 12:05:32
Job time : 223 secs

This Page Blank (uspto)

GenCore version 5.1.6

Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 03:18:00 ; Search time 49.1206 Seconds
(without alignments)
10851.002 Million cell updates/sec

Title: US-09-983-000A-1_COPY_1262_1272

Perfect score: 11

Sequence: 1 taactatatga 11

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.hcg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB ID	Description
C	1	11	100.0	26	8	ATH53625 Arabidops
C	2	11	100.0	60	1	AF292645 Versinia
C	3	11	100.0	63	8	ATH552136 Arabidops
C	4	11	100.0	113	6	CQ657839 Sequence
C	5	11	100.0	116	11	G37949 C4M51 Plasm
C	6	11	100.0	120	11	CR380577 Arabidops
C	7	11	100.0	124	3	AF113904 Entamoeba
C	8	11	100.0	131	8	AJ835129 Arabidops
C	9	11	100.0	131	8	AV201015 Arabidops
C	10	11	100.0	141	8	AX202071 Arabidops
C	11	11	100.0	158	6	AX899869 Sequence
C	12	11	100.0	158	6	BD035402 Sequence
C	13	11	100.0	162	8	ATH526739 Arabidops
C	14	11	100.0	162	8	ATH526751 Arabidops
C	15	11	100.0	163	3	APOCHORA2
C	16	11	100.0	166	3	AMU72269 A.polyphemu
C	17	11	100.0	166	3	AMU72270 Apis mellif
C	18	11	100.0	166	3	AMU72271 Apis mellif
C	19	11	100.0	168	8	CSP301708 Coniothyr

CQ659752 Sequence	180	6	CQ659752	11	100.0
AX202070 Arabidops	182	8	AX202070	11	100.0
AB092529 Sus scrofa	193	11	AB092529	11	100.0
BV151419 PZA02136-	196	11	BV151419	11	100.0
AF253885 Sus scrofa	197	11	AF253885	11	100.0
G03405 human STS W	208	11	G03405	11	100.0
AX916742 Sequence	220	6	AX916742	11	100.0
BD052275 Sequence	220	6	BD052275	11	100.0
U18058 Lycopersico	221	8	LEU18058	11	100.0
BV151414 PZA02136-	226	11	BV151414	11	100.0
AY427968 Dicerorhi	239	4	AY427968	11	100.0
AY606076 Dicerorhi	239	4	AY606076	11	100.0
CQ658479 Sequence	240	6	CQ658479	11	100.0
BV151421 PZA02136-	246	11	BV151421	11	100.0
BV151423 PZA02136-	246	11	BV151423	11	100.0
BV151412 PZA02136-	247	11	BV151412	11	100.0
AJ835161 Arabidops	251	8	AJ835161	11	100.0
AJ522030 Arabidops	251	8	ATH522030	11	100.0
AJ835175 Arabidops	253	8	AJ835175	11	100.0
BV151413 PZA02136-	253	11	BV151413	11	100.0
BV151416 PZA02136-	254	11	BV151416	11	100.0
BV151417 PZA02136-	254	11	BV151417	11	100.0
AX896745 Sequence	255	6	AX896745	11	100.0
BD032278 Sequence	255	6	BD032278	11	100.0
BV151415 PZA02136-	255	11	BV151415	11	100.0
BV151418 PZA02136-	255	11	BV151418	11	100.0
BV151420 PZA02136-	255	11	BV151420	11	100.0
BV151422 PZA02136-	256	11	BV151422	11	100.0
U32650 Hantavirus	258	14	HMU32650	11	100.0
CQ662259 Sequence	265	6	CQ662259	11	100.0
AL008888 H.sapiens	266	11	HSU86H4T3	11	100.0
AX196455 Sequence	268	6	AX196455	11	100.0
AF411617 Lutzomyia	270	3	AF411617	11	100.0
M83322 Lycopersico	271	3	TOMACCC	11	100.0
AF411622 Lutzomyia	276	3	AF411622	11	100.0
Z61356 H.sapiens C	278	9	HS50G6R	11	100.0
AR526418 Sequence	284	6	AR526418	11	100.0
AY300822 Homo sapi	290	9	AY300822	11	100.0
CQ685511 Sequence	307	6	CQ685511	11	100.0
AY543653 Meloidogy	308	3	AY543653	11	100.0
Z67617 H.sapiens D	309	11	HSB3222D9	11	100.0
Z67273 H.sapiens D	325	11	HSA107XB9	11	100.0
Z51967 H.sapiens (332	11	HSA085YC9	11	100.0
AR318868 Sequence	333	6	AR318868	11	100.0
AJ407089 Eretmodus	336	5	ECY407089	11	100.0
AJ407092 Eretmodus	336	5	ECY407092	11	100.0
AR378076 Sequence	336	6	AR378076	11	100.0
CQ528347 Sequence	346	6	CQ528347	11	100.0
AF142541 Dasyuttil	347	3	AF142541	11	100.0
AX246514 Sequence	347	6	AX246514	11	100.0
CQ464327 Sequence	348	6	CQ464327	11	100.0
AR375988 Sequence	348	9	HS211WA9	11	100.0
Z23603 H.sapiens	348	9	HS211WA9	11	100.0
AY733638 Versinia	350	1	AY733638	11	100.0
AR395484 Sequence	351	6	AR395484	11	100.0
AF099961 Hirudo me	352	3	AF099961	11	100.0
G29668 human STS S	359	11	G29668	11	100.0
CQ485374 Sequence	370	6	CQ485374	11	100.0
AY15820 human STS C	375	11	G15820	11	100.0
AY134818 Lipeurus	380	3	AY134818	11	100.0
CQ669920 Sequence	385	6	CQ669920	11	100.0
AY733707 Macaca mu	386	9	AY733707	11	100.0
CQ711440 Sequence	387	6	CQ711440	11	100.0
AL00797 Human chrom	390	11	HUM4STS327	11	100.0
AL606884 Arabidops	391	11	AL606884	11	100.0
AX186547 Sequence	395	6	AX186547	11	100.0
AR249668 Sequence	397	6	AR249668	11	100.0
CQ663939 Sequence	400	6	CQ663939	11	100.0
AJ836680 Arabidops	401	8	AJ836680	11	100.0
CQ688193 Sequence	402	6	CQ688193	11	100.0
G72950 MARC 2666-2	404	11	G72950	11	100.0
L30557 Human STS U	409	11	HUMUT6377	11	100.0

c 93 11 100.0 416 8 ATH531472 AJ531472 Arabidops
 94 11 100.0 418 6 CQ517862 CQ517862 Sequence
 95 11 100.0 419 6 CQ694491 CQ694491 Sequence
 96 11 100.0 422 6 AR416058 AR416058 Sequence
 97 11 100.0 422 6 AX972892 AX972892 Sequence
 98 11 100.0 422 6 BD111611 BD111611 EST and e
 c 99 11 100.0 426 8 AY034293 AY034293 Wolffia a
 c 100 11 100.0 430 8 AJ836672 AJ836672 Arabidops

ALIGNMENTS

RESULT 1
 ATH53625/c
 LOCUS
 DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone
 363A10. linear PLN 29-MAR-2003
 ACCESSION AJ553625
 VERSION AJ553625.1 GI:29369788
 KEYWORDS left border; T-DNA flanking sequence.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1
 AUTHORS Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.,
 Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G.,
 Lepoint, L., Caboche, M. and Lecharny, A.
 T-DNA integration into the Arabidopsis genome depends on sequences

TITLE of pre-insertion sites
 JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
 MEDLINE 22363535
 PUBMED 12446565

REFERENCE 2 (bases 1 to 26)
 AUTHORS Balzergue, S.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
 Gaston Cremieux, 91057 Evry cedex, FRANCE
 COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana
 plants from INRA (Versailles). The DNA fragment (s) resulting from
 the PCR were directly sequenced from the left or the right border
 to determine the genomic sequence flanking the insertion. T-DNA
 derived sequences were removed. Information to order the
 corresponding mutant line and a link to a database providing a
 graphical display of the insertion site are available at
<http://dbgap.versailles.inra.fr/publiclines/>. This sequence has
 been generated in the framework of the French plant genomics
 program 'Genoplante' (<http://www.genoplante.com> and
<http://genoplante-info.infobiogen.fr>).
 Location/Qualifiers

FEATURES
 source
 1..26
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Wassiljewskijja"
 /db_xref="taxon:3702"
 /clone="363A10"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 misc_feature 1..26
 /note="T-DNA flanking sequence
 left border"

ORIGIN

Query Match 100.0%; Score 11; DB 8; Length 26;
 Best Local Similarity 100.0%; Pred. No. 2.7e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 17 TAACTATATGA 7

RESULT 2
 AF292645
 LOCUS
 DEFINITION Yersinia pestis Pla gene, partial cds.
 ACCESSION AF292645
 VERSION AF292645.1 GI:11177535
 KEYWORDS
 SOURCE Yersinia pestis
 ORGANISM Yersinia pestis

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Yersinia.

REFERENCE 1 (bases 1 to 60)
 AUTHORS Raoult, D., Aboudharam, G., Crubezy, E., Larrouy, G., Ludes, B. and
 Drancourt, M.

TITLE Molecular identification by 'suicide PCR' of Yersinia pestis as the
 agent of medieval black death

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (23), 12800-12803 (2000)
 MEDLINE 20524081
 PUBMED 11058154

REFERENCE 2 (bases 1 to 60)
 AUTHORS Raoult, D., Aboudharam, G. and Drancourt, M.
 TITLE Direct Submission

JOURNAL Submitted (02-AUG-2000) Unite des Rickettsiales, Faculte de Medecine,
 Boulevard Jean Moulin, Marseille 13385, France
 Location/Qualifiers

1..60
 /organism="Yersinia pestis"
 /mol_type="genomic DNA"
 /db_xref="taxon:632"
 /note="from individual with Medieval Black Death in 1348"
 complement(<1..>60)
 /codon_start=3
 /transl_table=11
 /product="Pla"
 /protein_id="AAG32324.1"
 /db_xref="GI:11177536"
 /translation="GGSYNNNGAYTGNPKGV"

ORIGIN

Query Match 100.0%; Score 11; DB 1; Length 60;
 Best Local Similarity 100.0%; Pred. No. 2.2e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 42 TAACTATATGA 52

RESULT 3
 ATH552136
 LOCUS

DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone
 329D10. linear PLN 29-MAR-2003

ACCESSION AJ552136
 VERSION AJ552136.1 GI:29368277
 KEYWORDS left border; T-DNA flanking sequence.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1
 AUTHORS Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.,
 Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G.,
 Lepoint, L., Caboche, M. and Lecharny, A.

TITLE T-DNA integration into the Arabidopsis genome depends on sequences
 of pre-insertion sites
 JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
 MEDLINE 22363535
 PUBMED 12446565

REFERENCE 2 (bases 1 to 63)
 AUTHORS Balzergue, S.
 TITLE Direct Submission

JOURNAL Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE

COMMENT PCR was performed on DNA from transformants of *Arabidopsis thaliana* plants from INRA (Versailles). The DNA fragment (g) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).

FEATURES

source
1..63
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Wassilewskija"
/db_xref="taxon:3702"
/clone="329D10"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
misc_feature
1..63
/note="T-DNA flanking sequence
left border"

ORIGIN

Query Match 100.0%; Score 11; DB 8; Length 63;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
|||||
Db 45 TAACTATATGCA 55

RESULT 4
CQ657839/c
LOCUS
DEFINITION Sequence 2765 from Patent WO02070737.
ACCESSION CQ657839
VERSION CQ657839.1 GI:42121700
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Pateht: WO 02070737-A 2765 12-SEP-2002;
Chondrogene Inc. (CA)

FEATURES

source
1..113
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 113;
Best Local Similarity 100.0%; Pred. No. 2e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
|||||
Db 15 TAACTATATGCA 5

RESULT 5
G37949
LOCUS
DEFINITION CAM51 Plasmodium falciparum haploid Plasmodium falciparum STS genomic, sequence tagged site.
ACCESSION G37949

VERSION G37949.1 GI:3056693
KEYWORDS STS.
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 116)
AUTHORS Su,X.-Z., Ferdig,M.T. and Wellens,T.E.
TITLE Integrated genetic and physical maps of Plasmodium falciparum
JOURNAL Unpublished (1998)
COMMENT

Contact: Thomas E. Wellens

LPD

NIAID

Bldg. 4, Room 126, NIH Campus, Bethesda, MD 20892-0425

Tel: (301) 496-4021

Fax: (301) 402-0079

Email: tew@helix.nih.gov

Primer A: ATGTAACATATATGAAATGTTTC

Primer B: AATAGGATTCATGCTAACCC

STS size: 116

PCR Profile:

Initial heat: 2min at 94°C

Cycles: 30 cycles, each having steps 20sec at 94°C;

10sec at 45°C; 10sec at 40°C; 30sec at 60°C

Machine: Perkin Elmer GeneAmp 9600

Protocol:

Template: 1-10 ng

Primer: 5 pM each

dNTPs: 0.3 ul of 10 mM stock solution

Taq Polymerase: 0.5 unit

Total Volume: 15 ul

Buffer:

10 mM Tris HCl, pH 8.3

1.5 mM MgCl2

50 mM KCl.

Location/Qualifiers

1..116

/organism="Plasmodium falciparum"

/mol_type="genomic DNA"

/strain="HB3XDD2"

/db_xref="taxon:5833"

/map="4"

/clone_lib="Plasmodium falciparum haploid"

/dev_stage="erythrocytic haploid stage"

1..116

primer_bind

1..21

primer_bind complement(96..116)

ORIGIN

Query Match 100.0%; Score 11; DB 11; Length 116;

Best Local Similarity 100.0%; Pred. No. 1.9e+04;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11

|||||

Db 4 TAACTATATGCA 14

RESULT 6

CR380577/c

LOCUS

DEFINITION

CR380577

ACCESSION

CR380577.1

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1

CR380577

Arabidopsis thaliana transposon insertion STS GT_5.114364, sequence

tagged site.

CR380577

CR380577.1

GI:45752041

STS; STS, sequence tagged site.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1

AUTHORS Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCulleggh, B., Walsh, S.,
JOURNAL Langham, S., LeGrys, C., Jones, J.D.G. and Bevan, M.
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 120)
TITLE Clarke, J.H.
JOURNAL Direct Submission
 Submitted (22-MAR-2004) Clarke J.H., John Innes Centre, Colney
 Lane, Norwich, NR4 7UJ, UK
COMMENT AT denotes an activation tag dissociation transposon within a
 single line, ET an enhancer trap dissociation transposon, GT a gene
 trap dissociation transposon, MT a mis-expression enhancer trap
 dissociation transposon, SM a defective suppressor mutator
 transposon. _3 denotes a sequence derived from the 3' end of the
 transposon. _5 denotes a sequence derived from the 5' end of the
 transposon. BBSRC GARNET, ATIS project
 On-line seed stock requests: <http://nasc.nott.ac.uk/> NASC stock
 code: N178239.

FEATURES Location/Qualifiers
 source
 1..120
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /variety="Landsberg erecta NASC stock code NW20"
 /cultivar="Unknown"
 /db_xref="taxon:3702"
 /clone="AC016827"
 1..120
 /standard_name="GT_5.114364"

STS
 1..120
 /standard_name="GT_5.114364"

ORIGIN
 Query Match 100.0%; Score 11; DB 11; Length 120;
 Best Local Similarity 100.0%; Pred. No. 1.9e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
 |||||
Db 16 TAACTATATGCA 6

RESULT 7
AF113904/c
LOCUS Entamoeba histolytica retroviral integration site flanking region
DEFINITION sequence.
ACCESSION AF113904
VERSION AF113904.1 GI:4731928
KEYWORDS .
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica
REFERENCE 1 (bases 1 to 124)
AUTHORS Eukaryota; Entamoebidae; Entamoeba.
 Que, X., Kim, D., Alagon, A., Hirata, K., Shike, H., Shimizu, C.,
 Gonzalez, A., Burns, J.C. and Reed, S.L.
TITLE Pantropic retroviral vectors mediate gene transfer and expression
 in Entamoeba histolytica
JOURNAL Mol. Biochem. Parasitol. 99 (2), 237-245 (1999)
MEDLINE 98270312
PUBMED 10340487
REFERENCE 2 (bases 1 to 124)
AUTHORS Que, X., Kim, D. and Reed, S.L.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1998) Pathology, UCSD Medical Center, 200 W.
 Arbor Drive, #8416, San Diego, CA 92103-8416, USA

FEATURES Location/Qualifiers
 source
 1..124
 /organism="Entamoeba histolytica"
 /mol_type="genomic DNA"
 /strain="HMI:IMSS"
 /db_xref="taxon:5759"
 1..124
 /note="retroviral integration site flanking region"

misc_feature
 1..124
 /note="retroviral integration site flanking region"

ORIGIN
 Query Match 100.0%; Score 11; DB 3; Length 124;
 Best Local Similarity 100.0%; Pred. No. 1.9e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 1.9e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
 |||||
Db 36 TAACTATATGCA 26

RESULT 8
AJ835129
LOCUS Arabidopsis thaliana T-DNA flanking sequence, left border, clone
DEFINITION 096G08.
ACCESSION AJ835129
VERSION AJ835129.1 GI:52539335
KEYWORDS left border; T-DNA flanking sequence.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1
REFERENCE Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.,
 Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G.,
 Lepiniec, L., Caboche, M. and Lecharny, A.
AUTHORS T-DNA integration into the Arabidopsis genome depends on sequences
 of pre-insertion sites
TITLE EMBO Rep. 3 (12), 1152-1157 (2002)
JOURNAL 22363535
MEDLINE 12446565
PUBMED 12446565
REFERENCE 2 (bases 1 to 131)
AUTHORS Balzergue, S.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2004) Balzergue S., UMRGV, INRA/CNRS, 2 rue
 Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana
 plants from INRA (Versailles). The DNA fragment(s) resulting from
 the PCR were directly sequenced from the left or the right border
 to determine the genomic sequence flanking the insertion. T-DNA
 derived sequences were removed. Information to order the
 corresponding mutant line and a link to a database providing a
 graphical display of the insertion site are available at
<http://dbgap.versailles.inra.fr/publiclines/>. This sequence has
 been generated in the framework of the French plant genomics
 program 'Genoplante' (<http://www.genoplante.com> and
<http://genoplante-info.infobiogen.fr>).
FEATURES Location/Qualifiers
 source
 1..131
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Wassilewskija"
 /db_xref="taxon:3702"
 /clone="096G08"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 /ecotype="Wassilewskija"
 1..131
 /note="T-DNA flanking sequence
 left border"

ORIGIN
 Query Match 100.0%; Score 11; DB 8; Length 131;
 Best Local Similarity 100.0%; Pred. No. 1.9e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
 |||||
Db 26 TAACTATATGCA 36

RESULT 9
AY201015
LOCUS Arabidopsis thaliana sequence flanking Ds3 end of Ds-Trap insertion
DEFINITION

```

from line GT10014.
ACCESSION AY201015
VERSION AY201015.1 GI:27896969
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 131)
May, B.P., Smorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R.,
McCombie, W.R. and Martienssen, R.A.
Arabidopsis genomic sequences flanking Ds enhancer and gene traps
in transgenic lines
2 (bases 1 to 131)
May, B.P., Smorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R.,
McCombie, W.R. and Martienssen, R.A.
Direct Submission
Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown
Road, Cold Spring Harbor, NY 11724, USA
http://genetrap.cshl.org.
FEATURES
    source
        1..131
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /db_xref="taxon:3702"
            /clone="GT10014.Ds3.08.30.2001.jw09.131"
            /ecotype="Landsberg"
            /note="transgenic line GT10014"

ORIGIN
Query Match 100.0%; Score 11; DB 8; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.9e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 62 TAACTATATGA 72

RESULT 10
AY202071/c
LOCUS AY202071 141 bp DNA linear PLN 24-JAN-2003
DEFINITION Arabidopsis thaliana sequence flanking Ds3 end of Ds-Trap insertion
from line GT6062.
ACCESSION AY202071
VERSION AY202071.1 GI:27898025
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 141)
May, B.P., Smorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R.,
McCombie, W.R. and Martienssen, R.A.
Arabidopsis genomic sequences flanking Ds enhancer and gene traps
in transgenic lines
Unpublished
2 (bases 1 to 141)
May, B.P., Smorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R.,
McCombie, W.R. and Martienssen, R.A.
Direct Submission
Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown
Road, Cold Spring Harbor, NY 11724, USA
http://genetrap.cshl.org.
FEATURES
    source
        1..141
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /db_xref="taxon:3702"
            /clone="GT6062.Ds3.11.01.99.b.141"

from line GT10014.
/ecotype="Landsberg"
/note="transgenic line GT6062"

ORIGIN
Query Match 100.0%; Score 11; DB 8; Length 141;
Best Local Similarity 100.0%; Pred. No. 1.9e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 76 TAACTATATGA 66

RESULT 11
AX899869
LOCUS AX899869 158 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 15732 from Patent EP1033401.
ACCESSION AX899869
VERSION AX899869.1 GI:40054783
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Dumas Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.
TITLES Expressed sequence tags and encoded human proteins
JOURNAL Patent: EP 1033401-A 15732 06-SEP-2000;
Genset (FR)
FEATURES
    Location/Qualifiers
        source
            1..158
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 11; DB 6; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.8e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 27 TAACTATATGA 37

RESULT 12
BD035402
LOCUS BD035402 158 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD035402
VERSION BD035402.1 GI:22577144
KEYWORDS JP 2001269182-A/11648.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 158)
Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.
Arabidopsis genomic sequences flanking Ds enhancer and gene traps
in transgenic lines
Unpublished
2 (bases 1 to 141)
May, B.P., Smorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R.,
McCombie, W.R. and Martienssen, R.A.
Direct Submission
Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown
Road, Cold Spring Harbor, NY 11724, USA
http://genetrap.cshl.org.
FEATURES
    source
        1..141
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /db_xref="taxon:3702"
            /clone="GT6062.Ds3.11.01.99.b.141"

```

FEATURES	source	Location/Qualifiers	Location/Qualifiers
	1. .158		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	100.0%;	Score 11; DB 6; Length 158;	
Best Local Similarity	100.0%;	Pred. NO. 1.8e+04;	
Matches 11;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;	
Qy	1	TAACATATATGA 11	
Db	27	TAACATATATGA 37	
RESULT 13			
ATH526739/c			
LOCUS	ATH526739	162 bp	DNA linear PLN 29-MAR-2003
DEFINITION	Arabidopsis thaliana T-DNA flanking sequence, left border, clone 125F05.		
ACCESSION	AJ526739		
VERSION	AJ526739.1	GI:26794999	
KEYWORDS	left border; T-DNA flanking sequence.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.		
REFERENCE	1		
AUTHORS	Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G., Lepiniec, L., Caboche, M. and Lecharny, A.		
TITLE	T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites		
JOURNAL	EMBO Rep. 3 (12), 1152-1157 (2002)		
MEDLINE	22363535		
PUBMED	12446565		
REFERENCE	2 (bases 1 to 162)		
AUTHORS	Balzerque, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE		
COMMENT	PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at http://dbgap.versailles.inra.fr/publiclines/ . This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (http://www.genoplante.com and http://genoplante-info.infobiogen.fr).		
FEATURES	source	Location/Qualifiers	
	1. .162		
	/organism="Arabidopsis thaliana"		
	/mol_type="genomic DNA"		
	/cultivar="Wassillewskija"		
	/db_xref="taxon:3702"		
	/clone="125G11"		
	/clone_lib="Arabidopsis thaliana T-DNA insertion lines"		
	misc_feature		
	1. .162		
	/note="T-DNA flanking sequence left border"		
ORIGIN			
Query Match	100.0%;	Score 11; DB 8; Length 162;	
Best Local Similarity	100.0%;	Pred. NO. 1.8e+04;	
Matches 11;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;	
Qy	1	TAACATATATGA 11	
Db	84	TAACATATATGA 74	
RESULT 15			
APOCHORA2/c			
LOCUS	APOCHORA2	163 bp	DNA linear INV 26-APR-1993
DEFINITION	A.polyphemus (silkmoth) chorion gene 401a, 5' flank.		
ACCESSION	M10462		
VERSION	M10462.1	GI:155808	
KEYWORDS	chorion protein.		
SEGMENT	2 of 2		
SOURCE	Antheraea polyphemus (polyphemus moth)		

ORGANISM Antheraea polyphemus
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Saturniidae; Saturniinae; Saturniini; Antheraea.
REFERENCE 1 (bases 1 to 163)
AUTHORS Jones, C.W. and Kafatos, F.C.
TITLE Structure, organization and evolution of developmentally regulated
chorion genes in a silkworm
JOURNAL Cell 22 (3), 855-867 (1980)
MEDLINE 81112154
PUBMED 7460015
COMMENT Original source text: A.polyphemus (silkworm) DNA, clone Apc110.
FEATURES
source
Location/Qualifiers
1..163
/organism="Antheraea polyphemus"
/mol_type="genomic DNA"
/db_xref="taxon:7120"
ORIGIN About .1 kb away from segment 1.
Query Match 100.0%; Score 11; DB 3; Length 163;
Best Local Similarity 100.0%; Pred. No. 1.8e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAACTATATGCA 11
Db 29 TAACTATATGCA 19
RESULT 16
AMU72269/c
LOCUS 166 bp DNA linear INV 23-JAN-1997
DEFINITION Apis mellifera cytochrome oxidase subunit III (COIII) gene,
mitochondrial gene encoding mitochondrial protein, partial cds.
ACCESSION U72269
VERSION 1
KEYWORDS
SOURCE mitochondrion Apis mellifera (honey bee)
ORGANISM Apis mellifera
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
Apidae; Apis.
REFERENCE 1 (bases 1 to 166)
AUTHORS Koulianos, S. and Crozier, R.H.
TITLE Mitochondrial sequence characterisation of Australian commercial
and feral honeybee strains (Hymenoptera: Apidae; Apis mellifera
Linnaeus), in the context of the species worldwide
JOURNAL J. Australian Entomol. Soc. (1997) In press
REFERENCE 2 (bases 1 to 166)
AUTHORS Koulianos, S.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-1996) Experimental Ecology, ETH Zurich, ETH
Zentrum NW, Zurich 8092, Switzerland
FEATURES
source
Location/Qualifiers
1..166
/organism="Apis mellifera"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/db_xref="taxon:7460"
/haplotype="1"
gene
CDS
1..166
/gene="COIII"
/codon_start=1
/transl_table=5
/product="cytochrome oxidase subunit III"
/protein_id="AAB41167.1"
/db_xref="GI:1786007"
/translation="MKKNPFPFHMVTNPPWPIILSPFMTLISTIWIYSSISMPFMI
NFNSILIMWL"
ORIGIN
Query Match 100.0%; Score 11; DB 3; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.8e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAACTATATGCA 11
Db 31 TAACTATATGCA 21
RESULT 18
AMU72271/c
LOCUS 166 bp DNA linear INV 23-JAN-1997
DEFINITION Apis mellifera cytochrome oxidase subunit III (COIII) gene,
mitochondrial gene encoding mitochondrial protein, partial cds.
ACCESSION U72271
VERSION 1
KEYWORDS
SOURCE mitochondrion Apis mellifera (honey bee)
ORGANISM Apis mellifera
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;

Best Local Similarity 100.0%; Pred. No. 1.8e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAACTATATGCA 11
Db 31 TAACTATATGCA 21
RESULT 17
AMU72270/c
LOCUS 166 bp DNA linear INV 23-JAN-1997
DEFINITION Apis mellifera cytochrome oxidase subunit III (COIII) gene,
mitochondrial gene encoding mitochondrial protein, partial cds.
ACCESSION U72270
VERSION 1
KEYWORDS
SOURCE mitochondrion Apis mellifera (honey bee)
ORGANISM Apis mellifera
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
Apidae; Apis.
REFERENCE 1 (bases 1 to 166)
AUTHORS Koulianos, S. and Crozier, R.H.
TITLE Mitochondrial sequence characterisation of Australian commercial
and feral honeybee strains (Hymenoptera: Apidae; Apis mellifera
Linnaeus), in the context of the species worldwide
JOURNAL J. Australian Entomol. Soc. (1997) In press
REFERENCE 2 (bases 1 to 166)
AUTHORS Koulianos, S.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-1996) Experimental Ecology, ETH Zurich, ETH
Zentrum NW, Zurich 8092, Switzerland
FEATURES
source
Location/Qualifiers
1..166
/organism="Apis mellifera"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/db_xref="taxon:7460"
/haplotype="2"
gene
CDS
1..166
/gene="COIII"
/codon_start=1
/transl_table=5
/product="cytochrome oxidase subunit III"
/protein_id="AAB41168.1"
/db_xref="GI:1786009"
/translation="MKKNPFPFHMVTNPPWPIILSPFMTLISTIWIYSSISMPFMI
NFNSILIMWL"
ORIGIN
Query Match 100.0%; Score 11; DB 3; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.8e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAACTATATGCA 11
Db 31 TAACTATATGCA 21
RESULT 18
AMU72271/c
LOCUS 166 bp DNA linear INV 23-JAN-1997
DEFINITION Apis mellifera cytochrome oxidase subunit III (COIII) gene,
mitochondrial gene encoding mitochondrial protein, partial cds.
ACCESSION U72271
VERSION 1
KEYWORDS
SOURCE mitochondrion Apis mellifera (honey bee)
ORGANISM Apis mellifera
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;


```

source
1. .182
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="GT6062.Da3.03.03.00.b.182"
/ecotype="Landsberg"
/notes="transgenic line GT6062"

ORIGIN
Query Match      100.0%; Score 11; DB 8; Length 182;
Best Local Similarity 100.0%; Pred. No. 1.8e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 72 TAACTATATGA 62

RESULT 22
AB092529
LOCUS      AB092529      193 bp      DNA      linear      STS 02-AUG-2003
DEFINITION Sus scrofa DNA, chromosome 4, SJ678, sequence tagged site.
ACCESSION  AB092529
VERSION     AB092529.1 GI:33411581
KEYWORDS   STS.
SOURCE      Sus scrofa (pig)
ORGANISM    Sus scrofa
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1
AUTHORS     Fujishima-Kanaya,N., Toki,D., Suzuki,K., Sawazaki,T., Hiraiwa,H.,
            Iida,M., Hayashi,T., Uenishi,H., Wada,Y., Ito,Y. and Awata,T.
TITLE       Development of gene-associated markers using BAC clones and
            construction of a linkage map of swine chromosome 4
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 193)
AUTHORS     Fujishima-Kanaya,N., Toki,D., Suzuki,K., Sawazaki,T., Hiraiwa,H.,
            Iida,M., Hayashi,T., Uenishi,H., Wada,Y., Ito,Y. and Awata,T.
TITLE       Direct Submission
JOURNAL     Submitted (25-SEP-2002) Daisuke Toki, STAFF Institute, Animal
            Genome Analysis Team; 446-1 Ippaizuka, Kamiyokoba, Tsukuba, Ibaraki
            305-0854, Japan (E-mail:toki@gene.staff.or.jp, Tel:81-298-38-2190,
            Fax:81-298-38-2337)
COMMENT     Swine genomic BAC library in STAFF-Institute (available in NIAS DNA
            Bank, dna@genome.nias-k.affrc.go.jp).
FEATURES
source
1. .193
/organism="Sus scrofa"
/mol_type="genomic DNA"
/db_xref="taxon:9823"
/chromosome="4"
/clone="251A11"

primer_bind 1. .23
/PCR conditions="30 cycles 95degC 30 sec, 55degC 30 sec
            and 72degC 30 sec"
repeat_region 79. .86
/notes="microsatellite marker SJ678"
/rpt_type=tandem
/rpt_unit="ac"
repeat_region 89. .92
/notes="microsatellite marker SJ678"
/rpt_type=tandem
/rpt_unit="ac"
repeat_region 95. .110
/notes="microsatellite marker SJ678"
/rpt_type=tandem
/rpt_unit="ac"
primer_bind complement(170. .193)

ORIGIN
Query Match      100.0%; Score 11; DB 11; Length 193;
Best Local Similarity 100.0%; Pred. NO. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1. .196
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="NC7A"
/db_xref="taxon:4577"
/clone_lib="Zea mays NC7A"
/dev_stage="seedling"
/notes="Organ: leaf; genomic DNA from inbred line"
<1. .>196

Qy 1 TAACTATATGA 11
Db 159 TAACTATATGA 169

RESULT 23
BV151419
LOCUS      BV151419      196 bp      DNA      linear      STS 11-MAY-2004
DEFINITION PZA02136-70489-NC7A Zea mays NC7A Zea mays STS genomic, sequence
            tagged site.
ACCESSION  BV151419
VERSION     BV151419.1 GI:47100876
KEYWORDS   STS.
SOURCE      Zea mays
            Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 196)
AUTHORS     McMullen,M.D., Vroh Bi,I., Schultz,L., Duru,N., Schroeder,S.S.,
            Sanchez-Villeda,H., Houchins,K., Pang,Z., Polacco,M., Gardiner,J.
            and Coe,E.H.Jr.
TITLE       MNP SNP Discovery
JOURNAL     Unpublished (2003)
COMMENT     Contact: Michael D. McMullen
            Maize Mapping Project
            University of Missouri-Columbia
            301 Curtis Hall, Columbia, MO 65211-7020, USA
            Tel: 5738827606
            Fax: 5738847850
            Email: McMullenM@missouri.edu
            Primer A: CTGCTCTTCCGACCTTCTCTTA
            Primer B: TCAGACCGTTTCATTCGACAGT
            Protocol:
            PCR amplification of genomic DNA
            Template: 50 ng
                    each 0.5 uM
            dNTPs: each 200 uM
            Taq Polymerase: Redtaq (Sigma)
            Total Vol: 10 ul
            Amplicon sequencing
            ABI protocol - using d-Rhodamine terminator cycle
            sequencing ready reaction with ampliraq DNA polymerase FS
            Sequence ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
Redtaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)

Phred/Phrap quality scores: 15 18 12 8 8 13 19 29 37 37 42 47 47
47 47 37 37 37 37 47 68 68 68 57 47 42 35 35 33 42 36 37 40
40 40 40 48 50 48 43 46 50 50 50 68 68 57 57 57 57 57 57
57 51 50 57 57 57 57 57 57 57 54 54 54 57 50 50 50 50
50 57 57 57 57 47 47 57 57 54 54 57 57 57 57 57 57
52 52 52 57 57 57 57 57 57 57 57 57 57 57 57 57 57 57
57 57 57 57 57 57 57 57 57 57 57 57 57 57 57 57 57 57
44 28 28 32 32 43 40 54 54 54 54 59 41 41 18 18 15 20 15 12
15 12 13 13 15 15

FEATURES
source
1. .196
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="NC7A"
/db_xref="taxon:4577"
/clone_lib="Zea mays NC7A"
/dev_stage="seedling"
/notes="Organ: leaf; genomic DNA from inbred line"
<1. .>196

STS

```

```

ORIGIN
Query Match      100.0%; Score 11; DB 11; Length 196;
Best Local Similarity 100.0%; Pred. No. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 105 TAACTATATGA 115

RESULT 24
AF253885
LOCUS      AF253885                197 bp    DNA        linear    STS 25-MAY-2000
DEFINITION Sus scrofa chromosome 3 clone SWR1521, sequence tagged site.
ACCESSION  AF253885
VERSION     AF253885.1 GI:8072158
KEYWORDS   STS.
SOURCE     Sus scrofa (pig)
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1 (bases 1 to 197)
AUTHORS   Alexander,L.J., Rohrer,G.A. and Beattie,C.W.
TITLE     Cloning and characterization of 414 polymorphic porcine
            microsatellites
JOURNAL    Anim; Genet. 27 (3), 137-148 (1996)
MEDLINE    96314767
PUBMED     8759413
REFERENCE  2 (bases 1 to 197)
AUTHORS   Alexander,L.J., Rohrer,G.A. and Beattie,C.W.
TITLE     Direct Submission
JOURNAL    Submitted (06-APR-2000) USDA, ARS, US Meat Animal Research Center,
            P. O. Box 166, Clay Center, NE 68933-0166, USA
FEATURES   source
            1..197
            /organism="Sus scrofa"
            /mol_type="genomic DNA"
            /db_xref="taxon:9823"
            /chromosome="3"
            /clones="SWR1521"
            primer_bind
            primer_bind      complement(163..187)

ORIGIN
Query Match      100.0%; Score 11; DB 11; Length 197;
Best Local Similarity 100.0%; Pred. No. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 180 TAACTATATGA 190

RESULT 25
G03405/c
LOCUS      G03405                208 bp    DNA        linear    STS 19-OCT-1995
DEFINITION human STS WI-4855, sequence tagged site.
ACCESSION  G03405
VERSION     G03405.1 GI:720363
KEYWORDS   STS; STS sequence; primer; sequence tagged site.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 208)
AUTHORS   Hudson,T.
TITLE     Whitehead Institute/MIT Center for Genome Research; Random Genome
            Wide STS
JOURNAL    Unpublished (1995)
REFERENCE  2 (bases 1 to 208)
AUTHORS   Hudson,T.
TITLE     Whitehead Institute/MIT Center for Genome Research; Physically

```

```

Mapped STSs
Unpublished (1995)
GDB DSEG: D138884
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu

Primer A: TGGTATGTCATTTTGAATCC
Primer B: GGGAGAGCAAGTAAGGAAA
STS size: 151
PCR Profile:
    Presoak:
        Denaturation:
            Annealing: 56 degrees C
        Polymerization:
            PCR Cycles: 35
        Thermal Cycler:
    Protocol:
        Template: 10 ng
        Primer: each 5 pM
        dNTPs: each 4 nM
        Taq Polymerase: 0.025 units/ul
        Total Vol: 20 ul

Buffer:
    MgCl2: 1.5 mM
    KCl: 50 mM
    Tris-HCl: 10 mM
    pH: 9.3.
    Location/Qualifiers
        1..208
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /map="764_E.6; 766_A.1; 787_A.4; 810_G.1; 922_G.1;
        944_B.7; 749_B.11; 801_B.11; 882_C.12; 188.9 CR from top
        of Chr13 linkage group"
        55..205
        55..77
        complement(186..205)

STS
primer_bind
primer_bind      complement(186..205)

ORIGIN
Query Match      100.0%; Score 11; DB 11; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 97 TAACTATATGA 87

RESULT 26
AX916742
LOCUS      AX916742                220 bp    DNA        linear    PAT 18-DEC-2003
DEFINITION Sequence 32605 from Patent EP1033401.
ACCESSION  AX916742
VERSION     AX916742.1 GI:40210531
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE     Expressed sequence tags and encoded human proteins
JOURNAL    Patent: EP 1033401-A 32605 06-SEP-2000;
            Genset (FR)
FEATURES   Location/Qualifiers

```

```

source
1. .220
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 11; DB 6; Length 220;
Best Local Similarity 100.0%; Pred. No. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 201 TAACTATATGA 211

RESULT 27
BD052275
LOCUS BD052275 220 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD052275
VERSION BD052275.1 GI:22594017
KEYWORDS JP 2001269182-A/28521.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 220)
AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE Sequence tag and encoded human protein.
JOURNAL Patent: JP 2001269182-A 28521 02-OCT-2001;
GENSET
COMMENT
OS Homo sapiens (human)
PN JP 2001269182-A/28521
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN RAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
PI JORDAN
PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
C12N5/10,
PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
G06F15/40
CC
FH Key Location/Qualifiers.
FEATURES
source
1..220
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 11; DB 6; Length 220;
Best Local Similarity 100.0%; Pred. No. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 201 TAACTATATGA 211

RESULT 28
LEU18058/c
LOCUS LEU18058 221 bp DNA linear PLN 08-MAR-1997
DEFINITION Lycopersicon esculentum 1-aminocyclopropane-1-carboxylate synthase
(LB-ACSS) DNA, partial cdb.
ACCESSION U18058
VERSION U18058.1 GI:1561701
KEYWORDS
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE
1 (bases 1 to 221)
AUTHORS Kawakita,K., Hennig,L., Lincoln,J.E., Rottmann,W.H., Campbell,A.D.,
Zarembinski,T.I., Yu,G.X., Taylor,L.D. and Theologis,A.
TITLE The tomato 1-aminocyclopropane-1-carboxylate (ACC) synthase
multigene family is encoded by at least eight members
Unpublished
JOURNAL Theologis,A.
AUTHORS Direct Submission
TITLE Submitted (01-DEC-1994) Athanasios Theologis, Plant Gene Expression
Center, ARS/USDA, 800 Buchanan Street, Albany, CA 94710 and
JOURNAL Department of Plant Biology, University of California, Berkeley, CA
94720, USA
FEATURES
source
1..221
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/cultivar="Rutgers"
/db_xref="taxon:4081"
/clone="PSF31 (stock #394)"
1..221
/gene="LB-ACSS"
<1..>221
/gene="LB-ACSS"
/EC number="4.4.1.14"
/function="converts SAM to ACC; plant hormone
biosynthesis"
/standard_name="Lycopersicon esculentum 1-amino-
cyclopropane-1-carboxylate synthase"
/note="direct precursor of ethylene; one member of a
multigene family; from T2 region to the reaction center;
similar to the product encoded by PIR Accession Number
C41985"
/codon_start=1
/product="ACC synthase"
/protein_id="AAB48947.1"
/db_xref="GI:1561702"
/translation="ILITPAMTKNIHVSDIYAGTVFSDPKFVSIIALIDRKLEKS
KMNQVHIVSSLSKDLGLPGFRGMYSN"

ORIGIN
Query Match 100.0%; Score 11; DB 8; Length 221;
Best Local Similarity 100.0%; Pred. No. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 43 TAACTATATGA 33

RESULT 29
BV151414
LOCUS BV151414 226 bp DNA linear STS 11-MAY-2004
DEFINITION PZA02136-70475-GT119 Zea mays GT119 Zea mays STS genomic, sequence
tagged site.
ACCESSION BV151414
VERSION BV151414.1 GI:47100871
KEYWORDS STS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 226)
AUTHORS McMullen,M.D., Vroh Bi,I., Schultz,L., Duru,N., Schroeder,S.S.,
Sanchez-Villeda,H., Houchins,K., Fang,Z., Polacco,M., Gardiner,J.
and Coe,E.H.Jr.
TITLE MWP SNP Discovery
JOURNAL MWP SNP Discovery
COMMENT Unpublished (2003)
Contact: Michael D. McMullen
Maize Mapping Project
University of Missouri-Columbia

```

301 Curtis Hall, Columbia, MO 65211-7020, USA

Tel: 5739827606

Fax: 5739847850

Email: McMullen@missouri.edu

Primer A: CTGCTTCTCCGACCTTGCTA

Primer B: TCAAGACCTTTTCATTCGACGT

Protocol:

PCR amplification of genomic DNA

Template: 50 ng

Primer: each 0.5 uM

dNTPs: each 200 uM

Taq Polymerase: RedTaq (Sigma)

Total Vol: 10 uL

Amplicon sequencing

ABI protocol - using d-Rhodamine terminator cycle sequencing ready reaction with ampliTaq DNA polymerase FS
Sequence ran on ABI 3700 sequencer.

Buffer:

Genomic DNA amplification

RedTaq (Sigma)

Sequencing buffer

d-Rhodamine kit (ABI)

Phred/Phrap quality scores: 28 32 32 33 47 47 24 15 9 9 10 22 22
23 22 19 26 25 27 19 17 8 8 11 23 34 37 33 28 28 30 31 35 35 35
35 30 35 36 50 50 50 40 40 51 51 68 57 57 57 54 57 57 68 68 68
68 68 68 68 68 57 57 57 57 57 57 57 57 57 57 57 57 57 57 57
57 57 68 68 68 57 57 57 57 57 57 57 57 57 57 57 57 57 57 57
57 57 57 68 68 68 68 68 57 57 57 57 57 57 57 57 57 57 57 57
68 68 68 57 57 57 57 57 57 57 57 57 57 57 57 57 57 57 57 57
54 54 57 57 68 68 68 59 59 57 57 57 57 57 57 57 57 57 57 57
57 57 57 57 57 57 57 59 59 59 59 59 57 50 50 50 50 50 57 57
57 57 57 53 59 48 50 50 50 42 37 37 18.

FEATURES
source

1..226
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="Grl19"
/db_xref="taxon:4577"
/clone_lib="Zea mays Grl19"
/dev_stage="seedling"
/note="Organ: leaf; genomic DNA from inbred line"
<1..226

STS
ORIGIN

Query Match 100.0%; Score 11; DB 11; Length 226;
Best Local Similarity 100.0%; Pred. No. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
|||||
Db 123 TAACTATATGCA 133

RESULT 30
AY427968

LOCUS Dicerorhinus sumatrensis 239 bp DNA linear MAM 10-MAY-2004
DEFINITION Dicerorhinus sumatrensis microsatellite 191 sequence.
ACCESSION AY427968
VERSION AY427968.1 GI:37963468

KEYWORDS Dicerorhinus sumatrensis (Sumatran rhinoceros)
SOURCE Dicerorhinus sumatrensis
ORGANISM Dicerorhinus sumatrensis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Dicerorhinus.

1 (bases 1 to 239)
Scott, C., Foote, T., Morales, J.C., Fernando, P., Melnick, D.J.,
Boag, P.T., Davila, J.A. and van Coeverden de Groot, P.J.

TITLE Optimization of novel polymorphic microsatellites in the endangered
Sumatran rhinoceros (Dicerorhinus sumatrensis)
JOURNAL Mol. Ecol. Notes 4 (2), 194-196 (2004)

REFERENCE
AUTHORS

2 (bases 1 to 239)
Scott, C.A., Davila, J.A., Morales, J.C., Fernando, P., Melnick, D.,
Boag, P.T. and van Coeverden de Groot, P.J.

TITLE Direct Submission

JOURNAL Submitted (04-OCT-2003) Biology, Queen's University, BioSciences
Complex, Barrie Street, Kingston, Ontario K7L 3N6, Canada

FEATURES

source
1..239
/organism="Dicerorhinus sumatrensis"
/mol_type="genomic DNA"
/db_xref="taxon:89632"
/clone="SR191"
repeat_region
1..239
/note="microsatellite 191"
/rpt_type=tandem

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
|||||
Db 152 TAACTATATGCA 162

RESULT 31

AY606076

LOCUS Dicerorhinus sumatrensis 239 bp DNA linear MAM 07-JUN-2004
DEFINITION Dicerorhinus sumatrensis clone SuRh191 microsatellite sequence.

ACCESSION AY606076

VERSION AY606076.1 GI:47933795

KEYWORDS

SOURCE Dicerorhinus sumatrensis (Sumatran rhinoceros)
ORGANISM Dicerorhinus sumatrensis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Dicerorhinus.

REFERENCE
AUTHORS

1 (bases 1 to 239)
van Coeverden de Groot, P.J., Scott, C.A., D'Avila, J., Foote, T.,
Morales, J.-C., Melnick, D.J. and Boag, P.T.

TITLE Demographics of the endangered sumatran rhinoceros

JOURNAL Unpublished

REFERENCE

2 (bases 1 to 239)
van Coeverden de Groot, P.J., Scott, C.A., D'Avila, J., Foote, T.,
Morales, J.-C., Melnick, D.J. and Boag, P.T.

AUTHORS

Direct Submission

JOURNAL Submitted (25-APR-2004) Biology, Queen's University, BioSciences
Complex, Kingston, ON K7L 3N6, Canada

FEATURES

source
1..239
/organism="Dicerorhinus sumatrensis"
/mol_type="genomic DNA"
/db_xref="taxon:89632"
/clone="SuRh191"
repeat_region
1..239
/note="microsatellite"
/rpt_type=tandem

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
|||||
Db 152 TAACTATATGCA 162

RESULT 32

CQ658479/c

LOCUS Dicerorhinus sumatrensis 240 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 3405 from Patent WO02070737.

ACCESSION CQ658479

```

VERSION      CQ658479.1  GI:42124196
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS      Liew, C.C., Marshall, W.B. and Zhang, H.
TITLE        Compositions and methods relating to osteoarthritis
JOURNAL      Patent: WO 02070737-A 3405 12-SEP-2002;
              Chondrogene Inc. (CA)
FEATURES     Location/Qualifiers
              1..240
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 11; DB 6; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      161 TAACTATATGA 151

RESULT 33
BV151421
LOCUS      PZA02136-70488-Tx303 246 bp DNA linear STS 11-MAY-2004
DEFINITION tagged site.
VERSION     BV151421.1  GI:47100878
KEYWORDS    STS.
SOURCE      Zea mays
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
              1 (bases 1 to 246)
McMullen, M.D., Vroh Bi, I., Schultz, L., Duru, N., Schroeder, S.S.,
Sanchez-Villeda, H., Houchins, K., Fang, Z., Polacco, M., Gardiner, J.
and Coe, E.H.Jr.
MMP SNP Discovery
Unpublished (2003)

Contact: Michael D. McMullen
Maize Mapping Project
University of Missouri-Columbia
301 Curtis Hall, Columbia, MO 65211-7020, USA
Tel: 5738827606
Fax: 5738847850
Email: McMullen@missouri.edu
Primer A: CTGCTCTCCGACCTTGCTA
Primer B: TCAAGACCGTTTCATCTGACAGT
Protocol:
PCR amplification of genomic DNA
Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: RedTaq (Sigma)
Total Vol: 10 ul
Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
sequencing ready reaction with amplitaq DNA polymerase FS
Sequence ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
RedTaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)

VERSION      CQ658479.1  GI:42124196
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS      Liew, C.C., Marshall, W.B. and Zhang, H.
TITLE        Compositions and methods relating to osteoarthritis
JOURNAL      Patent: WO 02070737-A 3405 12-SEP-2002;
              Chondrogene Inc. (CA)
FEATURES     Location/Qualifiers
              1..240
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 11; DB 6; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      161 TAACTATATGA 151

RESULT 33
BV151421
LOCUS      PZA02136-70488-Tx303 246 bp DNA linear STS 11-MAY-2004
DEFINITION tagged site.
VERSION     BV151421.1  GI:47100878
KEYWORDS    STS.
SOURCE      Zea mays
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
              1 (bases 1 to 246)
McMullen, M.D., Vroh Bi, I., Schultz, L., Duru, N., Schroeder, S.S.,
Sanchez-Villeda, H., Houchins, K., Fang, Z., Polacco, M., Gardiner, J.
and Coe, E.H.Jr.
MMP SNP Discovery
Unpublished (2003)

Contact: Michael D. McMullen
Maize Mapping Project
University of Missouri-Columbia
301 Curtis Hall, Columbia, MO 65211-7020, USA
Tel: 5738827606
Fax: 5738847850
Email: McMullen@missouri.edu
Primer A: CTGCTCTCCGACCTTGCTA
Primer B: TCAAGACCGTTTCATCTGACAGT
Protocol:
PCR amplification of genomic DNA
Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: RedTaq (Sigma)
Total Vol: 10 ul
Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
sequencing ready reaction with amplitaq DNA polymerase FS
Sequence ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
RedTaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)

Phred/Phrap quality scores: 15 22 28 41 41 48 47 47 39 39 42 42 42
42 48 48 50 50 57 57 57 50 50 50 68 57 81 79 80 76 74 74 67 63
63 63 48 54 57 57 54 69 89 90 90 87 88 90 90 90 90 90 90
90 90 72 72 57 57 47 62 72 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 72 72 50 2 48 48 62 90 90 90 90 90 90 90
90 90 90 90 90 86 83 80 84 85 63 63 54 54 54 57 57
72 72 85 79 89 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 81 78 79 79 84 78 90 87 87
85 85 90 90 90 90 75 75 81 75 77 74 74 83 89 90 90 90 90
90 90 90 90 90 88 69 72 57 57 54 50 50 48 48 50 54 57 57
57 57 59 59 68 59 57 50 50 50 48 47 57 57 57 57 43 57 59 59
48 48 48 48 33 24 19 12 17 15 15 0.

FEATURES     Location/Qualifiers
              1..246
               /organism="Zea mays"
               /mol_type="genomic DNA"
               /cultivar="Tx303"
               /db_xref="taxon:4577"
               /clone_lib="Zea mays Tx303"
               /dev_stage="seedling"
               /note="Organ: leaf; genomic DNA from inbred line"
               <1..246
               complement (223..246)

ORIGIN
Query Match      100.0%; Score 11; DB 11; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      138 TAACTATATGA 148

RESULT 34
BV151423
LOCUS      PZA02136-70490-W22 R-scm2 246 bp DNA linear STS 11-MAY-2004
DEFINITION sequence tagged site.
VERSION     BV151423.1  GI:47100880
KEYWORDS    STS.
SOURCE      Zea mays
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
              1 (bases 1 to 246)
McMullen, M.D., Vroh Bi, I., Schultz, L., Duru, N., Schroeder, S.S.,
Sanchez-Villeda, H., Houchins, K., Fang, Z., Polacco, M., Gardiner, J.
and Coe, E.H.Jr.
MMP SNP Discovery
Unpublished (2003)

Contact: Michael D. McMullen
Maize Mapping Project
University of Missouri-Columbia
301 Curtis Hall, Columbia, MO 65211-7020, USA
Tel: 5738827606
Fax: 5738847850
Email: McMullen@missouri.edu
Primer A: CTGCTCTCCGACCTTGCTA
Primer B: TCAAGACCGTTTCATCTGACAGT
Protocol:
PCR amplification of genomic DNA
Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: RedTaq (Sigma)
Total Vol: 10 ul
Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
sequencing ready reaction with amplitaq DNA polymerase FS
Sequence ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
RedTaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)

```

sequencing ready reaction with amplitaq DNA polymerase FS
Sequence ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
RedTaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)

Phred/Phrap quality scores: 15 22 39 37 42 37 37 37 37 47 47 39 39
59 57 57 57 57 57 50 50 50 53 53 57 74 81 82 76 72 74 85 71 79
79 76 84 75 90 87 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 89 90 88 90 90 81 81 89 90 90 90 90 90 90
90
90
90
90 90 90 90 90 90 80 80 80 80 72 72 72 80 80 90 90 90 90 90
90 90 90 90 88 88 72 72 69 57 57 57 57 54 54 54 54 54 54
54 59 59 59 59 59 59 57 57 57 47 47 50 57 57 57 48 48 57 59 53
50 48 42 22 0 0 0 0 0 0 0.

FEATURES

source

1. .246
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="W22_R-scm2"
/db_xref="taxon:4577"
/clone_lib="Zea mays W22_R-scm2"
/dev_stage="seedling"
/note="Organ: leaf; genomic DNA from inbred line"
<1. .246

STS

ORIGIN

Query Match 100.0%; Score 11; DB 11; Length 246;
Best Local Similarity 100.0%; Pred. NO. 1.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11

Db 137 TAACTATATGCA 147

RESULT 35

BV151412

LOCUS PZA0136-70474-B73 Zea mays B73 Zea mays STS genomic, sequence
DEFINITION tagged site.

ACCESSION BV151412

VERSION BV151412.1 GI:47100869

KEYWORDS STS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 247)

McMullen, M.D., Vroh Bi, I., Schultz, L., Duru, N., Schroeder, S.S.,
Sanchez-Villeda, H., Houchins, K., Fang, Z., Polacco, M., Gardiner, J.
and Coe, E.H., Jr.

MMP SNP Discovery

Unpublished (2003)

REFERENCE

AUTHORS

Contact: Michael D. McMullen
Maize Mapping Project
University of Missouri-Columbia
301 Curtis Hall, Columbia, MO 65211-7020, USA
Tel: 5738827606

Fax: 5738847850

Email: McMullenM@missouri.edu

Primer A: CTGCTCTCCCGACCTTGCTA

Primer B: TCAAGACCGTTTCATTCGACGT

Protocol:

PCR amplification of genomic DNA

Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: RedTaq (Sigma)
Total Vol: 10 uL

Amplicon sequencing

ABI protocol - using d-Rhodamine terminator cycle
sequencing ready reaction with amplitaq DNA polymerase FS
Sequencing ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
RedTaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)

Phred/Phrap quality scores: 22 25 39 30 32 16 16 17 17 25 30 36 36
39 39 42 48 48 50 53 57 59 57 57 79 70 71 67 76 76 58 58
43 43 48 48 57 57 72 72 90 90 90 90 90 90 89 89 90 90 90
90 90 90 90 90 90 76 76 70 72 72 65 60 60 63 90 85 78 88 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 88 87 86 80 76 83 72 72 80 87 90 90 90 90
90 90 90 90 88 88 69 69 57 57 57 57 57 59 59 59 59 59
59 68 68 68 68 50 50 50 50 50 50 50 50 50 50 50 50 50 50
59 48 48 48 41 28 19 0 0 0 0 0.

FEATURES

source

1. .247
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="Zea mays B73"
/dev_stage="seedling"
/note="Organ: leaf; genomic DNA from inbred line"
<1. .246
complément(223. .246)

STS

ORIGIN

Query Match 100.0%; Score 11; DB 11; Length 247;

Best Local Similarity 100.0%; Pred. NO. 1.6e+04;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11

Db 137 TAACTATATGCA 147

RESULT 36

AJ835161/c

LOCUS

DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone
099A06.

ACCESSION AJ835161

VERSION AJ835161.1 GI:52539367

KEYWORDS left border; T-DNA flanking sequence.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

AUTHORS

Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.,
Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G.,
Lepiniec, L., Caboche, M. and Lecharny, A.

T-DNA integration into the Arabidopsis genome depends on sequences
of pre-insertion sites

EMBO Rep. 3 (12), 1152-1157 (2002)

JOURNAL

MEDLINE

22363535

FUBMED

12446565

REFERENCE

2 (bases 1 to 251)

AUTHORS Balzergue, S.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2004) Balzergue S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT PCR was performed on DNA from transformants of *Arabidopsis thaliana* plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).

FEATURES source 1..251
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Wassillewskija"
 /db_xref="taxon:3702"
 /clone="099A06"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 /ecotype="Wassillewskija"
 misc_feature 1..251
 /note="T-DNA flanking sequence
 left border"

ORIGIN
 Query Match 100.0%; Score 11; DB 8; Length 251;
 Best Local Similarity 100.0%; Pred. No. 1.6e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAACTATATGA 11
 |||||
 Db 75 TAACTATATGA 65

RESULT 37
 ATH522030/c 251 bp DNA linear PLN 29-MAR-2003
LOCUS Arabidopsis thaliana T-DNA flanking sequence, left border, clone 283H05.
DEFINITION AJ522030.1 GI:26790266
VERSION left border; T-DNA flanking sequence.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, P., Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G., Lepiniec, L., Caboche, M. and Lecharny, A.
TITLE T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites
JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
MEDLINE 22363535
PUBMED 12446565
REFERENCE 2 (bases 1 to 251)
AUTHORS Balzergue, S.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT PCR was performed on DNA from transformants of *Arabidopsis thaliana* plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbgap.versailles.inra.fr/publiclines/>. This sequence has

been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).

FEATURES source 1..251
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Wassillewskija"
 /db_xref="taxon:3702"
 /clone="283H05"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 /note="T-DNA flanking sequence
 left border"

ORIGIN
 Query Match 100.0%; Score 11; DB 8; Length 251;
 Best Local Similarity 100.0%; Pred. No. 1.6e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAACTATATGA 11
 |||||
 Db 70 TAACTATATGA 60

RESULT 38
 AJ835175/c 253 bp DNA linear PLN 22-SEP-2004
LOCUS Arabidopsis thaliana T-DNA flanking sequence, left border, clone 099E05.
DEFINITION AJ835175.1 GI:52539381
VERSION left border; T-DNA flanking sequence.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, P., Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G., Lepiniec, L., Caboche, M. and Lecharny, A.
TITLE T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites
JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
MEDLINE 22363535
PUBMED 12446565
REFERENCE 2 (bases 1 to 253)
AUTHORS Balzergue, S.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2004) Balzergue S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT PCR was performed on DNA from transformants of *Arabidopsis thaliana* plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).

FEATURES source 1..253
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Wassillewskija"
 /db_xref="taxon:3702"
 /clone="099E05"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 /ecotype="Wassillewskija"
 misc_feature 1..253

```

/notes="T-DNA flanking sequence
left border"

ORIGIN
Query Match      100.0%; Score 11; DB 8; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
    |||||
Db 77 TAACTATATGA 67

RESULT 39
BV151413
LOCUS PZA02136-70476-CO159 Zea mays CO159 Zea mays STS genomic, sequence
DEFINITION tagged site.
ACCESSION BV151413
VERSION BV151413.1 GI:47100870
KEYWORDS STS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 253)
McMullen,M.D., Vroh Bi,I., Schultz,L., Duru,N., Schroeder,S.S.,
Sanchez-Villeda,H., Houchins,K., Fang,Z., Polacco,M., Gardiner,J.
MMP SNP Discovery
and Coe,E.H.Jr.
Unpublished (2003)

Contact: Michael D. McMullen
Maize Mapping Project
University of Missouri-Columbia
301 Curtis Hall, Columbia, MO 65211-7020, USA
Tel: 5738827606
Fax: 5738847850
Email: McMullenM@missouri.edu
Primer A: CTGCTCTTCCGACCTTGCTA
Primer B: TCAAGACCGTTTCATTCTGACAGT
Protocol:
PCR amplification of genomic DNA
Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: RedTaq (Sigma)
Total Vol: 10 ul
Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
sequencing ready reaction with amplitaq DNA polymerase FS
Sequence ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
RedTaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)

Phred/Phrap quality scores: 0 15 16 20 20 24 37 33 48 59 59 48 48
48 48 48 43 43 43 43 43 43 48 50 57 57 57 57 57 57 59 59 59 59
74 72 76 76 78 63 63 63 63 63 63 63 63 63 63 63 63 63 63 63
90 90 89 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 89 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
69 69 72 57 57 57 57 59 59 68 68 68 68 57 57 57 50 48 48 57
57 57 57 57 57 59 59 57 48 48 48 33 33 19 17 10 19 19 17 0.

FEATURES
Location/Qualifiers

```

```

source 1. .253
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="CO159"
/db_xref="taxon:4577"
/clone_lib="Zea mays CO159"
/dev_stage="seedling"
/notes="Organ: leaf; genomic DNA from inbred line"
1. .253
primer_bind 1. .22
primer_bind complement (230. .253)
ORIGIN
Query Match      100.0%; Score 11; DB 11; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
    |||||
Db 144 TAACTATATGA 154

RESULT 40
BV151416
LOCUS PZA02136-70491-ILO Zea mays ILO Zea mays STS genomic, sequence
DEFINITION tagged site.
ACCESSION BV151416
VERSION BV151416.1 GI:47100873
KEYWORDS STS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 254)
McMullen,M.D., Vroh Bi,I., Schultz,L., Duru,N., Schroeder,S.S.,
Sanchez-Villeda,H., Houchins,K., Fang,Z., Polacco,M., Gardiner,J.
MMP SNP Discovery
and Coe,E.H.Jr.
Unpublished (2003)

Contact: Michael D. McMullen
Maize Mapping Project
University of Missouri-Columbia
301 Curtis Hall, Columbia, MO 65211-7020, USA
Tel: 5738827606
Fax: 5738847850
Email: McMullenM@missouri.edu
Primer A: CTGCTCTTCCGACCTTGCTA
Primer B: TCAAGACCGTTTCATTCTGACAGT
Protocol:
PCR amplification of genomic DNA
Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: RedTaq (Sigma)
Total Vol: 10 ul
Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
sequencing ready reaction with amplitaq DNA polymerase FS
Sequence ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
RedTaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)

Phred/Phrap quality scores: 0 0 0 0 0 0 15 19 26 39 39 40 43 43
43 43 43 40 40 39 35 35 36 40 52 58 75 67 57 64 55 59 49 51
49 45 47 49 40 47 41 33 35 45 49 47 52 54 68 58 71 66 76 73 75 74
80 79 87 87 85 76 85 52 50 35 25 27 19 19 24 30 30 35 28 43 51 47

```



```

82 83 71 87 81 74 76 78 86 90 84 83 90 86 90 89 90 81 79 79 90
90 90 90 90 90 90 90 90 90 88 73 76 63 77 75 86 79 78
71 67 70 67 86 86 90 82 82 87 75 84 70 75 72 88 90 90 89 80 88 74
74 74 84 84 79 76 81 81 90 90 83 87 81 90 77 77 90 76 73
70 68 77 81 70 80 80 63 63 42 42 35 30 30 36 51 56 80 85 83 90
89 85 90 90 90 86 90 87 89 82 76 63 63 48 48 43 48 48 42 42
39 35 37 47 43 54 54 48 59 59 42 37 30 30 30 37 37 48 39 47
30 28 14 24 24 37 37 41 41 37 37 37 19 0 0 0 0 0 0.

FEATURES
  source
    Location/Qualifiers
      1..254
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /cultivar="ILO"
        /db_xref="taxon:4577"
        /clone_lib="Zea mays ILO"
        /dev_stage="seedling"
        /note="Organ: leaf; genomic DNA from inbred line"
        <1..>254

STS
  source
    Location/Qualifiers
      1..254
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /cultivar="Mol7"
        /db_xref="taxon:4577"
        /clone_lib="Zea mays Mol7"
        /dev_stage="seedling"
        /note="Organ: leaf; genomic DNA from inbred line"
        <1..>254

ORIGIN
  Query Match 100.0%; Score 11; DB 11; Length 254;
  Best Local Similarity 100.0%; Pred. No. 1.6e+04;
  Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
  |||||
Db 145 TAACTATATGA 155

RESULT 41
BV151417
LOCUS 254 bp DNA linear STS 11-MAY-2004
DEFINITION PZA02136-70486-Mol7 Zea mays Mol7 Zea mays STS genomic, sequence
tagged site.
ACCESSION BV151417
VERSION BV151417.1 GI:47100874
KEYWORDS STS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 254)
McMullen,M.D., Vroh Bi,I., Schultz,L., Duru,N., Schroeder,S.S.,
Sanchez-Villeda,H., Houchins,K., Fang,Z., Polacco,M., Gardiner,J.
and Coe,E.H.Jr.
WMP SNP Discovery
Unpublished (2003)

Contact: Michael D. McMullen
Maize Mapping Project
University of Missouri-Columbia
301 Curtis Hall, Columbia, MO 65211-7020, USA
Tel.: 5738827606
Fax.: 5738847850
Email: McMullenM@missouri.edu
Primer A: CTGCTCTCCGACCTTGCTA
Primer B: TCAAGACCTTTCATTCGACAGT
Protocol:
PCR amplification of genomic DNA
Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: RedTaq (Sigma)
Total Vol: 10 ul
Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
sequencing ready reaction with ampliTaq DNA polymerase PS
Sequence ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
RedTaq (Sigma)

```

```

Sequencing buffer
d-Rhodamine kit (ABI)

Phred/Phrap quality scores: 0 0 0 0 0 0 0 18 32 32 28 20 19 19
30 25 23 23 18 18 13 13 13 24 26 39 41 20 20 20 28 24 26 13 14
14 19 24 17 17 24 24 30 37 30 39 30 32 30 48 40 63 61 64 72
63 61 61 69 69 74 61 45 40 27 17 19 11 26 26 34 42 47 44 49 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
79 63 70 70 84 72 72 63 48 48 57 57 57 57 47 47 57 57 57 54
54 54 57 57 57 57 57 57 57 57 59 59 59 57 57 57 47 47 57
47 47 48 48 47 47 47 57 59 59 59 57 57 54 54 54 54 57
46 43 39 37 32 32 42 48 48 48 59 43 41 30 19 14 14 28 26
25 22 24 17 15 15 12 12 15 28 32 35 35 35 39 32 18 16 0 0 0 0

FEATURES
  source
    Location/Qualifiers
      1..254
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /cultivar="Mol7"
        /db_xref="taxon:4577"
        /clone_lib="Zea mays Mol7"
        /dev_stage="seedling"
        /note="Organ: leaf; genomic DNA from inbred line"
        <1..>254

STS
  source
    Location/Qualifiers
      1..254
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /cultivar="Mol7"
        /db_xref="taxon:4577"
        /clone_lib="Zea mays Mol7"
        /dev_stage="seedling"
        /note="Organ: leaf; genomic DNA from inbred line"
        <1..>254

ORIGIN
  Query Match 100.0%; Score 11; DB 11; Length 254;
  Best Local Similarity 100.0%; Pred. No. 1.6e+04;
  Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
  |||||
Db 145 TAACTATATGA 155

RESULT 42
AX896745
LOCUS 255 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 12608 from Patent EP1033401.
ACCESSION AX896745
VERSION AX896745.1 GI:40051629
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE Expressed sequence tags and encoded human proteins
JOURNAL Patent: EP 1033401-A 12608 06-SEP-2000;
Genet (FR)

FEATURES
  source
    Location/Qualifiers
      1..255
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"

ORIGIN
  Query Match 100.0%; Score 11; DB 6; Length 255;
  Best Local Similarity 100.0%; Pred. No. 1.6e+04;
  Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
  |||||
Db 196 TAACTATATGA 206

RESULT 43
BD032278
LOCUS 255 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD032278

```

```
dNTPs:          each 200 uM
Taq Polymerase: Redtaq (Sigma)
Total Vol:      10 ul
Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
sequencing ready reaction with amplitaq DNA polymerase FS
Sequence ran on ABI 3700 sequencer.
```

```
Buffer:
Genomic DNA amplification
Redtaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)
```

```
Phred/Phrap quality scores: 0 0 0 0 0 0 0 0 19 34 41 47 47 41 48
48 43 43 36 36 43 43 48 72 72 72 88 70 63 74 74 81 70 71
72 74 76 75 55 49 33 37 42 43 43 58 58 76 75 82 90 87 89 90
90 90 90 90 90 90 90 90 90 85 87 79 71 74 74 87 81 81 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 87 90 90 90 90 90 90 90 90 90 90 79 74 81 90 90 90
90 90 87 90 90 90 90 90 90 90 89 84 84 84 90 90 90 87 90
90 84 79 73 73 80 90 80 90 85 85 85 90 90 90 82 90 80 83 88
87 90 79 73 71 54 54 39 18 12 11 13 11 18 56 56 54 52 54 65
71 89 90 90 90 90 90 90 90 85 76 85 63 63 48 57 48 48 62 62
63 54 57 54 54 54 54 48 59 59 59 59 57 48 48 42 42 47 57 57
48 48 48 47 47 47 41 41 25 25 0 0 0 0 0 0 0 0 0 0 .
```

```
FEATURES             Location/Qualifiers
     source            .
     . . . . .         /organism="Zea mays"
     . . . . .         /mol_type="genomic DNA"
     . . . . .         /cultivar="IHO"
     . . . . .         /db_xref="taxon:4577"
     . . . . .         /clone_lib="Zea mays IHO"
     . . . . .         /dev_stage="seedling"
     note="Organ: leaf; genomic DNA from inbred line"
     <1..>.255
```

```
STS
ORIGIN

Query Match       100.0%;   Score 11;   DB 11;   Length 255;
Best Local Similarity 100.0%;   Pred. No. 1.6e+04;
Matches 11; Conservative    0; Mismatches    0; Indels    0; Gaps    0;
```

```
Qy  1 TAACTATATGGA 11
      |||||
Db   145 TAACTATATGGA 155
```

```
RESULT 45
BV151418
LOCUS BV151418
DEFINITION PZA02136-70477-Mp708 Zea mays Mp708 Zea mays STS genomic, sequence tagged site.
ACCESSION BV151418
VERSION BV151418.1 GI:47100875
KEYWORDS STS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 255)
McMullen,M.D., Vroh Bi,I., Schultz,L., Duru,N., Schroeder S.S.,
Sanchez-Villeda,H., Houchins,K., Fang,Z., Polacco,M., Gardiner,J.,
and Coe,E.H.Jr.
MMP SNP Discovery
Unpublished (2003)
```

```
REFERENCE
AUTHORS McMullen,M.D., Vroh Bi,I., Schultz,L., Duru,N., Schroeder S.S.,
Sanchez-Villeda,H., Houchins,K., Fang,Z., Polacco,M., Gardiner,J.,
and Coe,E.H.Jr.
TITLE MMP SNP Discovery
JOURNAL Unpublished (2003)
COMMENT Contact: Michael D. McMullen
Maize Mapping Project
University of Missouri-Columbia
301 Curtis Hall, Columbia, MO 65211-7020, USA
Tel: 5738827606
Fax: 5738847850
```



```

QY 1 TAACTATATGA 11
Db 199 TAACTATATGA 189

RESULT 50
HSU86H4T3 266 bp DNA linear STS 19-MAY-1998
LOCUS H.sapiens STS from genomic clone U86H4, sequence tagged site.
DEFINITION AL008888
ACCESSION AL008888
VERSION 1 GI:2624174
KEYWORDS STS; single read.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 266)
AUTHORS Ross, M., Dunham, A., Howell, G., Steingrub, H., Coffey, A., McDowell, J., Whittaker, A., Warry, G., Joseph, S., Brady, N., Huckle, S., Taylor, R., Baron, L., Smith, A., Morrone, A., Kelly, P., Fowler, J., and Hunt, S.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1997) E-mail contact: humquery@sanger.ac.uk
COMMENT Marker atSG14965 (Primer A: TATGGGAGTCTTGGAAACG; Primer B: TGTTTGAGCCATACACAGG; amplicon size: 107 bp) is from sequence generated from the T3 end of COSMID U86H4. U86H4 is part of the bacterial clone contigs constructed by the Chromosome X Mapping Group. (http://www.sanger.ac.uk/HGP/ChrX/) U86H4 is from the Lawrence Livermore National Laboratory flow-sorted X chromosome cosmid library LLOXNC01.
FEATURES
    source
        Location/Qualifiers
            1..266
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="X"
                /clone="U86H4"

ORIGIN
Query Match 100.0%; Score 11; DB 11; Length 266;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
Db 136 TAACTATATGA 146

RESULT 51
AX196455 268 bp DNA linear PAT 07-SEP-2001
LOCUS Sequence 162 from Patent WO0151627.
DEFINITION AX196455
ACCESSION AX196455
VERSION 1 GI:15386661
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Glycine max
REFERENCE 1
AUTHORS Haug, B.M., Wang, M.L., Parsons, J.D. and Parnell, L.D.
TITLE Nucleic acid molecules and other molecules associated with soybean cyst nematode resistance
JOURNAL Patent: WO 0151627-A 162 19-JUL-2001;
MONGANTO COMPANY (US)
FEATURES
    source
        Location/Qualifiers
            1..268
                /organism="Glycine max"
                /mol_type="unassigned DNA"
                /db_xref="taxon:3847"
                /note="Seq ID: 240017 region G3 4566 16"

QY 1 TAACTATATGA 11
Db 199 TAACTATATGA 189

RESULT 50
HSU86H4T3 266 bp DNA linear STS 19-MAY-1998
LOCUS H.sapiens STS from genomic clone U86H4, sequence tagged site.
DEFINITION AL008888
ACCESSION AL008888
VERSION 1 GI:2624174
KEYWORDS STS; single read.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 266)
AUTHORS Ross, M., Dunham, A., Howell, G., Steingrub, H., Coffey, A., McDowell, J., Whittaker, A., Warry, G., Joseph, S., Brady, N., Huckle, S., Taylor, R., Baron, L., Smith, A., Morrone, A., Kelly, P., Fowler, J., and Hunt, S.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1997) E-mail contact: humquery@sanger.ac.uk
COMMENT Marker atSG14965 (Primer A: TATGGGAGTCTTGGAAACG; Primer B: TGTTTGAGCCATACACAGG; amplicon size: 107 bp) is from sequence generated from the T3 end of COSMID U86H4. U86H4 is part of the bacterial clone contigs constructed by the Chromosome X Mapping Group. (http://www.sanger.ac.uk/HGP/ChrX/) U86H4 is from the Lawrence Livermore National Laboratory flow-sorted X chromosome cosmid library LLOXNC01.
FEATURES
    source
        Location/Qualifiers
            1..266
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="X"
                /clone="U86H4"

ORIGIN
Query Match 100.0%; Score 11; DB 11; Length 266;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
Db 136 TAACTATATGA 146

RESULT 51
AX196455 268 bp DNA linear PAT 07-SEP-2001
LOCUS Sequence 162 from Patent WO0151627.
DEFINITION AX196455
ACCESSION AX196455
VERSION 1 GI:15386661
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Glycine max
REFERENCE 1
AUTHORS Haug, B.M., Wang, M.L., Parsons, J.D. and Parnell, L.D.
TITLE Nucleic acid molecules and other molecules associated with soybean cyst nematode resistance
JOURNAL Patent: WO 0151627-A 162 19-JUL-2001;
MONGANTO COMPANY (US)
FEATURES
    source
        Location/Qualifiers
            1..268
                /organism="Glycine max"
                /mol_type="unassigned DNA"
                /db_xref="taxon:3847"
                /note="Seq ID: 240017 region G3 4566 16"

QY 1 TAACTATATGA 11
Db 199 TAACTATATGA 189

RESULT 50
HSU86H4T3 266 bp DNA linear STS 19-MAY-1998
LOCUS H.sapiens STS from genomic clone U86H4, sequence tagged site.
DEFINITION AL008888
ACCESSION AL008888
VERSION 1 GI:2624174
KEYWORDS STS; single read.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 266)
AUTHORS Ross, M., Dunham, A., Howell, G., Steingrub, H., Coffey, A., McDowell, J., Whittaker, A., Warry, G., Joseph, S., Brady, N., Huckle, S., Taylor, R., Baron, L., Smith, A., Morrone, A., Kelly, P., Fowler, J., and Hunt, S.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1997) E-mail contact: humquery@sanger.ac.uk
COMMENT Marker atSG14965 (Primer A: TATGGGAGTCTTGGAAACG; Primer B: TGTTTGAGCCATACACAGG; amplicon size: 107 bp) is from sequence generated from the T3 end of COSMID U86H4. U86H4 is part of the bacterial clone contigs constructed by the Chromosome X Mapping Group. (http://www.sanger.ac.uk/HGP/ChrX/) U86H4 is from the Lawrence Livermore National Laboratory flow-sorted X chromosome cosmid library LLOXNC01.
FEATURES
    source
        Location/Qualifiers
            1..266
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="X"
                /clone="U86H4"

ORIGIN
Query Match 100.0%; Score 11; DB 11; Length 266;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
Db 136 TAACTATATGA 146

RESULT 51
AX196455 268 bp DNA linear PAT 07-SEP-2001
LOCUS Sequence 162 from Patent WO0151627.
DEFINITION AX196455
ACCESSION AX196455
VERSION 1 GI:15386661
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Glycine max
REFERENCE 1
AUTHORS Haug, B.M., Wang, M.L., Parsons, J.D. and Parnell, L.D.
TITLE Nucleic acid molecules and other molecules associated with soybean cyst nematode resistance
JOURNAL Patent: WO 0151627-A 162 19-JUL-2001;
MONGANTO COMPANY (US)
FEATURES
    source
        Location/Qualifiers
            1..268
                /organism="Glycine max"
                /mol_type="unassigned DNA"
                /db_xref="taxon:3847"
                /note="Seq ID: 240017 region G3 4566 16"

QY 1 TAACTATATGA 11
Db 199 TAACTATATGA 189

RESULT 50
HSU86H4T3 266 bp DNA linear STS 19-MAY-1998
LOCUS H.sapiens STS from genomic clone U86H4, sequence tagged site.
DEFINITION AL008888
ACCESSION AL008888
VERSION 1 GI:2624174
KEYWORDS STS; single read.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 266)
AUTHORS Ross, M., Dunham, A., Howell, G., Steingrub, H., Coffey, A., McDowell, J., Whittaker, A., Warry, G., Joseph, S., Brady, N., Huckle, S., Taylor, R., Baron, L., Smith, A., Morrone, A., Kelly, P., Fowler, J., and Hunt, S.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1997) E-mail contact: humquery@sanger.ac.uk
COMMENT Marker atSG14965 (Primer A: TATGGGAGTCTTGGAAACG; Primer B: TGTTTGAGCCATACACAGG; amplicon size: 107 bp) is from sequence generated from the T3 end of COSMID U86H4. U86H4 is part of the bacterial clone contigs constructed by the Chromosome X Mapping Group. (http://www.sanger.ac.uk/HGP/ChrX/) U86H4 is from the Lawrence Livermore National Laboratory flow-sorted X chromosome cosmid library LLOXNC01.
FEATURES
    source
        Location/Qualifiers
            1..266
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="X"
                /clone="U86H4"

ORIGIN
Query Match 100.0%; Score 11; DB 11; Length 266;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
Db 136 TAACTATATGA 146

RESULT 51
AX196455 268 bp DNA linear PAT 07-SEP-2001
LOCUS Sequence 162 from Patent WO0151627.
DEFINITION AX196455
ACCESSION AX196455
VERSION 1 GI:15386661
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Glycine max
REFERENCE 1
AUTHORS Haug, B.M., Wang, M.L., Parsons, J.D. and Parnell, L.D.
TITLE Nucleic acid molecules and other molecules associated with soybean cyst nematode resistance
JOURNAL Patent: WO 0151627-A 162 19-JUL-2001;
MONGANTO COMPANY (US)
FEATURES
    source
        Location/Qualifiers
            1..268
                /organism="Glycine max"
                /mol_type="unassigned DNA"
                /db_xref="taxon:3847"
                /note="Seq ID: 240017 region G3 4566 16"

QY 1 TAACTATATGA 11
Db 199 TAACTATATGA 189

RESULT 50
HSU86H4T3 266 bp DNA linear STS 19-MAY-1998
LOCUS H.sapiens STS from genomic clone U86H4, sequence tagged site.
DEFINITION AL008888
ACCESSION AL008888
VERSION 1 GI:2624174
KEYWORDS STS; single read.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 266)
AUTHORS Ross, M., Dunham, A., Howell, G., Steingrub, H., Coffey, A., McDowell, J., Whittaker, A., Warry, G., Joseph, S., Brady, N., Huckle, S., Taylor, R., Baron, L., Smith, A., Morrone, A., Kelly, P., Fowler, J., and Hunt, S.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1997) E-mail contact: humquery@sanger.ac.uk
COMMENT Marker atSG14965 (Primer A: TATGGGAGTCTTGGAAACG; Primer B: TGTTTGAGCCATACACAGG; amplicon size: 107 bp) is from sequence generated from the T3 end of COSMID U86H4. U86H4 is part of the bacterial clone contigs constructed by the Chromosome X Mapping Group. (http://www.sanger.ac.uk/HGP/ChrX/) U86H4 is from the Lawrence Livermore National Laboratory flow-sorted X chromosome cosmid library LLOXNC01.
FEATURES
    source
        Location/Qualifiers
            1..266
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="X"
                /clone="U86H4"

ORIGIN
Query Match 100.0%; Score 11; DB 11; Length 266;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 
```

```

PUBMED 1549612
REFERENCE 2 (bases 1 to 271)
AUTHORS Oetiker,J.H., Olson,D.C., Shiu,O.Y. and Yang,S.F.
TITLE Differential induction of seven 1-aminocyclopropane-1-carboxylate
synthase genes by elicitor in suspension cultures of tomato
(Lycopersicon esculentum)
JOURNAL Plant Mol. Biol. 34 (2), 275-286 (1997)
MEDLINE 97351561
PUBMED 9207843
COMMENT On Jun 12, 1993 this sequence version replaced gi:170509.
FEATURES
    source
        Location/Qualifiers
            1..271
                /organism="Lycopersicon esculentum"
                /mol_type="mRNA"
                /db_xref="taxon:4081"
                /tissue_types="calyx"
            <1..>271
                /EC number="4.4.1.14"
                /note="homologue"
                /codon_start=2
                /product="1-aminocyclopropane-1-carboxylate synthase"
                /protein_id="AAC37507.1"
                /db_xref="GI:867992"
                /translations="MRNELNLIITFAMTKNIHIVDEIVAGTVFSPKFSIIIEALI
                DRKLEKSMNQVHIYSSLSKDLGFCFRVGMYSNNETLAAATK"
ORIGIN
Query Match 100.0%; Score 11; DB 8; Length 271;
Best Local Similarity 100.0%; Pred. NO. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 65 TAACTATATGA 55

RESULT 54
AF411622/c
LOCUS AF411622 276 bp DNA linear INV 14-JAN-2002
DEFINITION Lutzomyia longipalpis clone LfST6024 microsatellite sequence.
ACCESSION AF411622
VERSION AF411622.1 GI:18140901
KEYWORDS
SOURCE Lutzomyia longipalpis
ORGANISM Lutzomyia longipalpis
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Psychodoidea;
Psychodidae; Lutzomyia; Lutzomyia.
REFERENCE 1 (bases 1 to 276)
AUTHORS Watt,P.C., Noyes,H.A. and Kemp,S.J.
TITLE Polymorphic microsatellite loci in the sand fly Lutzomyia
longipalpis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 276)
AUTHORS Watt,P.C., Noyes,H.A. and Kemp,S.J.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-2001) School of Biological Sciences, University
of Liverpool, Crown Street, Liverpool, Merseyside L69 7ZD, UK
FEATURES
    source
        Location/Qualifiers
            1..276
                /organism="Lutzomyia longipalpis"
                /mol_type="genomic DNA"
                /db_xref="taxon:7200"
                /clone="LfST6024"
            1..276
                /notes="microsatellite"
                /rpt_family="tandem"
    repeat_region
        1..276
            100.0%; Score 11; DB 3; Length 276;
            Best Local Similarity 100.0%; Pred. NO. 1.6e+04;
            Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
Query Match 100.0%; Score 11; DB 3; Length 276;
Best Local Similarity 100.0%; Pred. NO. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 65 TAACTATATGA 55

```

```

Qy 1 TAACTATATGA 11
Db 101 TAACTATATGA 91

RESULT 55
HS50G6R/c
LOCUS HS50G6R 278 bp DNA linear PRI 22-OCT-1995
DEFINITION H.sapiens CpG island DNA genomic MseI fragment, clone 5096, reverse
read cp95096.rta.
ACCESSION Z61356
VERSION Z61356.1 GI:1033734
KEYWORDS CpG island; genomic MseI fragment.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE Purification of CpG islands using a methylated DNA binding column
JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070
PUBMED 8012384
REFERENCE 2 (bases 1 to 278)
AUTHORS MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
COMMENT Vector: pGEM-5zf(-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES
    source
        Location/Qualifiers
            1..278
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /clone="5096"
                /sex="male"
                /tissue_type="blood"
                /clone_lib="CGI-1"
                /dev_stage="adult"
ORIGIN
Query Match 100.0%; Score 11; DB 9; Length 278;
Best Local Similarity 100.0%; Pred. NO. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 106 TAACTATATGA 96

RESULT 56
AR526418
LOCUS AR526418 284 bp DNA linear PAT 22-SEP-2004
DEFINITION Sequence 31378 from patent US 6703491.
ACCESSION AR526418
VERSION AR526418.1 GI:52461906
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 284)
AUTHORS Homburger,S.A., Ebens,A.J. Jr., Erickson,C.S., Francis-Lang,H.L.,
Margolis,J.S., Reddy,B.P., Ruddy,D.A. and Buchman,A.R.
TITLE Drosophila sequences
JOURNAL Patent: US 6703491-A 31378 09-MAR-2004;
FEATURES
    source
        Location/Qualifiers
            1..284
                /organism="unknown"
                /mol_type="genomic DNA"

```

```

ORIGIN
Query Match      100.0%; Score 11; DB 6; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
    |||||
Db 259 TAACTATATGA 269

RESULT 57
AY300822
LOCUS      Homo sapiens immunoglobulin heavy chain variable region VH 3-53
DEFINITION
ACCESSION  AY300822
VERSION     AY300822.1 GI:31790069
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE  1 (bases 1 to 290)
AUTHORS   Nardini,E., Neri,F., Vicenzi,B., Poli,G., Capello,D., Gaidano,G.,
            Vitolo,U., Menard,S. and Balsari,A.
TITLE     Thymic function and immunoglobulin mutation genotype in B-cell
            Chronic Lymphocytic Leukemia patients
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 290)
AUTHORS   Nardini,E., Neri,F., Vicenzi,B., Poli,G., Capello,D., Gaidano,G.,
            Vitolo,U., Menard,S. and Balsari,A.
TITLE     Direct Submission
JOURNAL   Submitted (15-MAY-2003) Experimental Oncology, Istituto Nazionale
            Tumori, via Venezian 1, Milan 20133, Italy
FEATURES   source
            1..290
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /isolation_source="B-CLL patient (N 22)"
            /db_xref="taxon:9606"
            /chromosome="14"
            /map="14q32"
            /product="immunoglobulin heavy chain variable region VH"
            <1..>290
            3-53 family"
            /codon_start=3
            /product="immunoglobulin heavy chain variable region VH"
            3-53 family"
            /protein_id="AAP58389.1"
            /db_xref="GI:31790070"
            /translation="VQLVESGGGLVQPGGSLRLSCASGGFTVLTNYSWVRQAPGKGL
            QWVSALFYGGDTTADSVKGRFTISRDNSKNTLYLQNSRVSDTAVYCAT"

ORIGIN
Query Match      100.0%; Score 11; DB 9; Length 290;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
    |||||
Db 92 TAACTATATGA 102

RESULT 58
CQ685511
LOCUS      Homo sapiens (human)
DEFINITION  Sequence 30437 from Patent WO20070737.
ACCESSION  CQ685511
VERSION     CQ685511.1 GI:42212820
KEYWORDS   Homo sapiens (human)

ORIGIN
Query Match      100.0%; Score 11; DB 6; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
    |||||
Db 259 TAACTATATGA 269

RESULT 57
AY300822
LOCUS      Homo sapiens immunoglobulin heavy chain variable region VH 3-53
DEFINITION
ACCESSION  AY300822
VERSION     AY300822.1 GI:31790069
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE  1 (bases 1 to 290)
AUTHORS   Nardini,E., Neri,F., Vicenzi,B., Poli,G., Capello,D., Gaidano,G.,
            Vitolo,U., Menard,S. and Balsari,A.
TITLE     Thymic function and immunoglobulin mutation genotype in B-cell
            Chronic Lymphocytic Leukemia patients
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 290)
AUTHORS   Nardini,E., Neri,F., Vicenzi,B., Poli,G., Capello,D., Gaidano,G.,
            Vitolo,U., Menard,S. and Balsari,A.
TITLE     Direct Submission
JOURNAL   Submitted (15-MAY-2003) Experimental Oncology, Istituto Nazionale
            Tumori, via Venezian 1, Milan 20133, Italy
FEATURES   source
            1..290
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /isolation_source="B-CLL patient (N 22)"
            /db_xref="taxon:9606"
            /chromosome="14"
            /map="14q32"
            /product="immunoglobulin heavy chain variable region VH"
            <1..>290
            3-53 family"
            /codon_start=3
            /product="immunoglobulin heavy chain variable region VH"
            3-53 family"
            /protein_id="AAP58389.1"
            /db_xref="GI:31790070"
            /translation="VQLVESGGGLVQPGGSLRLSCASGGFTVLTNYSWVRQAPGKGL
            QWVSALFYGGDTTADSVKGRFTISRDNSKNTLYLQNSRVSDTAVYCAT"

ORIGIN
Query Match      100.0%; Score 11; DB 9; Length 290;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
    |||||
Db 92 TAACTATATGA 102

RESULT 58
CQ685511
LOCUS      Homo sapiens (human)
DEFINITION  Sequence 30437 from Patent WO20070737.
ACCESSION  CQ685511
VERSION     CQ685511.1 GI:42212820
KEYWORDS   Homo sapiens (human)

ORIGIN
Query Match      100.0%; Score 11; DB 6; Length 307;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
    |||||
Db 115 TAACTATATGA 125

RESULT 59
AY543653
LOCUS      Meloidogyne hapla putative secretory esophageal gland protein
DEFINITION (sec1) gene, partial cds.
ACCESSION  AY543653
VERSION     AY543653.1 GI:44894060
KEYWORDS   Meloidogyne hapla
SOURCE     Meloidogyne hapla
ORGANISM   Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
            Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
REFERENCE  1 (bases 1 to 308)
AUTHORS   Lee,S., Yaghoobi,J. and Williamson,V.M.
TITLE     A putative secretory esophageal gland protein (sec-1) gene in the
            root-knot nematode Meloidogyne hapla
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 308)
AUTHORS   Lee,S., Yaghoobi,J. and Williamson,V.M.
TITLE     Direct Submission
JOURNAL   Submitted (04-FEB-2004) Nematology, University of California, One
            Shields Avenue, Davis, CA 95616, USA
FEATURES   source
            1..308
            /organism="Meloidogyne hapla"
            /mol_type="genomic DNA"
            /db_xref="taxon:6305"
            <1..>308
            /genes="sec1"
            /join(<1..36,128..>228)
            /genes="sec1"
            /product="putative secretory esophageal gland protein"
            /codon_start=1
            /product="putative secretory esophageal gland protein"
            /protein_id="AA548571.1"
            /db_xref="GI:44894061"
            /translation="ENNQPKSTLDSLKLQLEEVSKIAIEVIERDAKAAQIERMLHEA
            KS"

ORIGIN
Query Match      100.0%; Score 11; DB 3; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
    |||||
Db 251 TAACTATATGA 261

```

```

RESULT 60
HSB3222D9          309 bp  DNA   linear   STS 04-NOV-1995
LOCUS              H.sapiens DNA segment containing (CA) repeat; clone AFMb322zd9;
DEFINITION         single read, sequence tagged site.
ACCESSION          Z67617
VERSION            Z67617.1 GI:1052400
KEYWORDS            STS; CA repeat; dinucleotide repeat; GT repeat; microsatellite DNA;
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE           1 (bases 1 to 309)
AUTHORS             Weissbach,J.
TITLE               Direct Submission
JOURNAL             Submitted (31-OCT-1995) Genethon, B.P. 60, 91002 Evry Cedex France.
E-mail: Jean.Weissenbach@genethon.fr
COMMENT             full automatic.
FEATURES             Location/Qualifiers
                     source
                     1..309
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /cell_line="CEPH 134702"
                     /clone_lib="genomic DNA"
                     /notes="cloning vector is M13mp18"

ORIGIN
Query Match          100.0%; Score 11; DB 11; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TAACTATATGCA 11
Db      135 TAACTATATGCA 145

RESULT 61
HSA107XB9
LOCUS              HSA107XB9
DEFINITION         H.sapiens DNA segment containing (CA) repeat; clone AFMa107XB9;
ACCESSION          Z67273
VERSION            Z67273.1 GI:1052056
KEYWORDS            STS; CA repeat; dinucleotide repeat; GT repeat; microsatellite DNA;
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE           1 (bases 1 to 325)
AUTHORS             Weissbach,J.
TITLE               Direct Submission
JOURNAL             Submitted (31-OCT-1995) Genethon, B.P. 60, 91002 Evry Cedex France.
E-mail: Jean.Weissenbach@genethon.fr
COMMENT             full automatic.
FEATURES             Location/Qualifiers
                     source
                     1..325
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /cell_line="CEPH 134702"
                     /clone_lib="genomic DNA"
                     /notes="cloning vector is M13mp18"

ORIGIN
Query Match          100.0%; Score 11; DB 11; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TAACTATATGCA 11
Db      135 TAACTATATGCA 145

RESULT 62
HSA085YC9
LOCUS              HSA085YC9
DEFINITION         H.sapiens (D11S4205) DNA segment containing (CA) repeat; clone
APMa085YC9; single read, sequence tagged site.
ACCESSION          Z51967
VERSION            Z51967.1 GI:1233267
KEYWORDS            STS; CA repeat; dinucleotide repeat; GT repeat; microsatellite DNA;
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE           1 (bases 1 to 332)
AUTHORS             Dib,C., Faure,S., Fizames,C., Samson,D., Drouot,N., Vignal,A.,
Millasseau,P., Marc,S., Hazan,J., Seboun,E., Lathrop,M., Gyapay,G.,
Morissette,J. and Weissenbach,J.
TITLE               A comprehensive genetic map of the human genome based on 5,264
microsatellites
JOURNAL             Nature 380 (6570), 152-154 (1996)
MEDLINE             96176476
PUBMED              8600387
REFERENCE           2 (bases 1 to 332)
AUTHORS             Weissenbach,J.
TITLE               Direct Submission
JOURNAL             Submitted (01-SEP-1995) Genethon, B.P. 60, 91002 Evry Cedex France.
E-mail: Jean.Weissenbach@genethon.fr
COMMENT             full automatic.
FEATURES             Location/Qualifiers
                     source
                     1..332
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /cell_line="CEPH 134702"
                     /clone_lib="genomic DNA"
                     /notes="cloning vector is M13mp18"

ORIGIN
Query Match          100.0%; Score 11; DB 11; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TAACTATATGCA 11
Db      108 TAACTATATGCA 118

RESULT 63
AR318868
LOCUS              AR318868
DEFINITION         Sequence 1418 from patent US 6562958.
ACCESSION          AR318868
VERSION            AR318868.1 GI:33699971
KEYWORDS            Unknown.
SOURCE              Unknown.
ORGANISM            Unclassified.
REFERENCE           1 (bases 1 to 333)
AUTHORS             Breton,G. and Bush,D.
TITLE               Nucleic acid and amino acid sequences relating to Acinetobacter
baumanni for diagnostics and therapeutics
JOURNAL             Patent: US 6562958-A 1418 13-MAY-2003;
FEATURES             Location/Qualifiers
                     source
                     1..333
                     /organism="unknown"

```

```

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TAACTATATGCA 11
Db      85  TAACTATATGCA 95

RESULT 62
HSA085YC9
LOCUS              HSA085YC9
DEFINITION         H.sapiens (D11S4205) DNA segment containing (CA) repeat; clone
APMa085YC9; single read, sequence tagged site.
ACCESSION          Z51967
VERSION            Z51967.1 GI:1233267
KEYWORDS            STS; CA repeat; dinucleotide repeat; GT repeat; microsatellite DNA;
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE           1 (bases 1 to 332)
AUTHORS             Dib,C., Faure,S., Fizames,C., Samson,D., Drouot,N., Vignal,A.,
Millasseau,P., Marc,S., Hazan,J., Seboun,E., Lathrop,M., Gyapay,G.,
Morissette,J. and Weissenbach,J.
TITLE               A comprehensive genetic map of the human genome based on 5,264
microsatellites
JOURNAL             Nature 380 (6570), 152-154 (1996)
MEDLINE             96176476
PUBMED              8600387
REFERENCE           2 (bases 1 to 332)
AUTHORS             Weissenbach,J.
TITLE               Direct Submission
JOURNAL             Submitted (01-SEP-1995) Genethon, B.P. 60, 91002 Evry Cedex France.
E-mail: Jean.Weissenbach@genethon.fr
COMMENT             full automatic.
FEATURES             Location/Qualifiers
                     source
                     1..332
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /cell_line="CEPH 134702"
                     /clone_lib="genomic DNA"
                     /notes="cloning vector is M13mp18"

ORIGIN
Query Match          100.0%; Score 11; DB 11; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TAACTATATGCA 11
Db      108 TAACTATATGCA 118

RESULT 63
AR318868
LOCUS              AR318868
DEFINITION         Sequence 1418 from patent US 6562958.
ACCESSION          AR318868
VERSION            AR318868.1 GI:33699971
KEYWORDS            Unknown.
SOURCE              Unknown.
ORGANISM            Unclassified.
REFERENCE           1 (bases 1 to 333)
AUTHORS             Breton,G. and Bush,D.
TITLE               Nucleic acid and amino acid sequences relating to Acinetobacter
baumanni for diagnostics and therapeutics
JOURNAL             Patent: US 6562958-A 1418 13-MAY-2003;
FEATURES             Location/Qualifiers
                     source
                     1..333
                     /organism="unknown"

```



```

/moi_type="genomic DNA"

ORIGIN
Query Match          100.0%; Score 11; DB 6; Length 333;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
    |||||
Db 168 TAACTATATGA 178

RESULT 64
ECY407089/c
LOCUS      ECY407089          336 bp      DNA      linear      VRT 26-APR-2001
DEFINITION Eretmodus cyanostictus mitochondrial D-loop control region,
haplotype 12.
ACCESSION  AJ407089
VERSION     AJ407089.1 GI:12802095
KEYWORDS   control region; D-loop.
SOURCE     mitochondrial Eretmodus cyanostictus (Tanganyika clon)
ORGANISM   Eretmodus cyanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes;
            Labroidel; Cichlidae; Eretmodus.
REFERENCE   1 (bases 1 to 336)
AUTHORS    Ruber, L., Meyer, A., Sturmhuber, C. and Verheyen, E.
TITLE      Population structure in two sympatric species of the Lake
            Tanganyika cichlid tribe Eretmodini: evidence for introgression
JOURNAL    Mol. Ecol. 10 (5), 1207-1225 (2001)
MEDLINE    21275939
PUBMED     11380878
REFERENCE   2 (bases 1 to 336)
AUTHORS    Ruber, L.
TITLE      Direct Submission
SUBMITTED  (18-OCT-2000) Ruber L., Zoology, Museo Nacional de
Ciencias Naturales, Jose Gutierrez Abascal 2, MADRID, 28006, SPAIN
FEATURES   source
            /organism="Eretmodus cyanostictus"
            /organelle="mitochondrion"
            /mol_type="genomic DNA"
            /db_xref="taxon:52954"
            /haplotype="12"
            /country="Democratic Republic of the Congo"
            1..336
            /note="control region"

D-loop
ORIGIN
Query Match          100.0%; Score 11; DB 5; Length 336;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
    |||||
Db 221 TAACTATATGA 211

RESULT 66
ECY407112/c
LOCUS      ECY407112          336 bp      DNA      linear      VRT 26-APR-2001
DEFINITION Eretmodus cyanostictus mitochondrial D-loop control region,
haplotype 35.
ACCESSION  AJ407112
VERSION     AJ407112.1 GI:12802118
KEYWORDS   control region; D-loop.
SOURCE     mitochondrial Eretmodus cyanostictus (Tanganyika clon)
ORGANISM   Eretmodus cyanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes;
            Labroidel; Cichlidae; Eretmodus.
REFERENCE   1 (bases 1 to 336)
AUTHORS    Ruber, L., Meyer, A., Sturmhuber, C. and Verheyen, E.
TITLE      Population structure in two sympatric species of the Lake
            Tanganyika cichlid tribe Eretmodini: evidence for introgression
JOURNAL    Mol. Ecol. 10 (5), 1207-1225 (2001)
MEDLINE    21275939
PUBMED     11380878
REFERENCE   2 (bases 1 to 336)
AUTHORS    Ruber, L.
TITLE      Direct Submission
SUBMITTED  (18-OCT-2000) Ruber L., Zoology, Museo Nacional de
Ciencias Naturales, Jose Gutierrez Abascal 2, MADRID, 28006, SPAIN
FEATURES   source
            /organism="Eretmodus cyanostictus"
            /organelle="mitochondrion"
            /mol_type="genomic DNA"
            /db_xref="taxon:52954"
            /haplotype="35"
            /country="Democratic Republic of the Congo"
            1..336
            /note="control region"

D-loop
ORIGIN
Query Match          100.0%; Score 11; DB 5; Length 336;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
    |||||
Db 221 TAACTATATGA 211

RESULT 65
ECY407092/c
LOCUS      ECY407092          336 bp      DNA      linear      VRT 26-APR-2001
DEFINITION Eretmodus cyanostictus mitochondrial D-loop control region,
haplotype 15.
ACCESSION  AJ407092
VERSION     AJ407092.1 GI:12802098
KEYWORDS   control region; D-loop.
SOURCE     mitochondrial Eretmodus cyanostictus (Tanganyika clon)
ORGANISM   Eretmodus cyanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes;
            Labroidel; Cichlidae; Eretmodus.

```

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 221 TAACTATATGA 211

RESULT 67
 AR378076
 LOCUS AR378076 336 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 3082 from patent US 6605709.
 ACCESSION AR378076
 VERSION AR378076.1 GI:40081258
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 336)
 AUTHORS Breton, G.L.
 TITLE Nucleic acid and amino acid sequences relating to *Proteus mirabilis*
 for diagnostics and therapeutics
 JOURNAL Patent: US 6605709-A 3082 12-AUG-2003;
 FEATURES Location/Qualifiers
 source 1..336
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 100.0%; Score 11; DB 6; Length 336;
 Best Local Similarity 100.0%; Pred. No. 1.5e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 237 TAACTATATGA 247

RESULT 68
 CQ528347/c
 LOCUS CQ528347 346 bp DNA linear PAT 30-JAN-2004
 DEFINITION Sequence 60214 from Patent WO0160860.
 ACCESSION CQ528347
 VERSION CQ528347.1 GI:41494611
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Schlögel, R., Endege, W.O. and Monahan, J.E.
 TITLE Genes differentially expressed in human prostate cancer and their use
 JOURNAL Patent: WO 0160860-A 60214 23-AUG-2001;
 Millennium Predictive Medicine, Inc. (US)
 FEATURES Location/Qualifiers
 source 1..346
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 100.0%; Score 11; DB 6; Length 346;
 Best Local Similarity 100.0%; Pred. No. 1.5e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 52 TAACTATATGA 42

RESULT 69
 AF142541/c
 LOCUS AF142541 347 bp DNA linear INV 21-FEB-2000

DEFINITION Dasyutilla sp. JC3 cytochrome oxidase I (COI) gene, partial cds;
 mitochondrial gene for mitochondrial product.
 ACCESSION AF142541
 VERSION AF142541.1 GI:7012598
 KEYWORDS
 SOURCE mitochondrial Dasyutilla sp. JC3
 ORGANISM Dasyutilla sp. JC3
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
 Vespoidea; Mutillidae; Sphaerophthalminae; Dasyutilla.
 REFERENCE 1 (bases 1 to 347)
 AUTHORS Carpenter, J.M. and Wheeler, W.C.
 TITLE Towards simultaneous analysis of morphological and molecular data
 in Hymenoptera
 JOURNAL Zool. Scr. 28 (1), 1-10 (1999)
 REFERENCE 2 (bases 1 to 347)
 AUTHORS Carpenter, J.M. and Wheeler, W.C.
 TITLE Direct Submission
 JOURNAL Submitted (13-APR-1999) Department of Invertebrates, American
 Museum of Natural History, Central Park West @79th St., New York,
 NY 10024, USA
 FEATURES Location/Qualifiers
 source 1..347
 /organism="Dasyutilla sp. JC3"
 /organelle="mitochondrion"
 /mol_type="genomic DNA"
 /isolate="JC3"
 /db_xref="taxon:103257"
 /gene="COI"
 complement(<1..>347)
 CDS complement(<1..>347)
 /gene="COI"
 /codon_start=3
 /product="cytochrome oxidase I"
 /protein_id="AAF35158.1"
 /db_xref="GI:7012599"
 /translation="MLFNNESGKKEFTGLMIYAMLTIGFLGFIWAHMETVGLDV
 DTRAYSTATMIITAIPTGIKVFSLATINGMKQKSTPTMTWSLGFVFLDTIGLGLGV
 LSNSSIXYSSWY"

ORIGIN
 Query Match 100.0%; Score 11; DB 3; Length 347;
 Best Local Similarity 100.0%; Pred. No. 1.5e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 101 TAACTATATGA 91

RESULT 70
 AX246514
 LOCUS AX246514 347 bp DNA linear PAT 28-SEP-2001
 DEFINITION Sequence 1444 from Patent WO0166753.
 ACCESSION AX246514
 VERSION AX246514.1 GI:15861188
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D.,
 Sudduth-Klinger, J., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,
 Kassam, A., Lamson, G., Drmanac, R., Crkvenjakov, R., Dickson, M.,
 Drmanac, S., Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and
 Stache-Crain, B.
 TITLE Human genes and gene expression products
 JOURNAL Patent: WO 0166753-A 1444 13-SEP-2001;
 Chiron Corporation (US) ; Hyseq Inc. (US)
 FEATURES Location/Qualifiers
 source 1..347

```

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 11; DB 6; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      104 TAACTATATGA 114

RESULT 71
LOCUS      CQ464327          348 bp      DNA      linear      PAT 30-JAN-2004
DEFINITION Sequence 7105 from Patent WO0192581.
ACCESSION  CQ464327
VERSION    CQ464327.1 GI:41429946
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Algate, P.A., Harlocker, S.L. and Jones, R.
TITLE      Compositions and methods for the therapy and diagnosis of
JOURNAL    ovariocancer
PATENT     WO 0192581-A 7105 06-DEC-2001;
CORIXA     CORPORATION (US)
FEATURES   Location/Qualifiers
            source
              1..348
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 11; DB 6; Length 348;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      241 TAACTATATGA 251

RESULT 72
LOCUS      AR375988          348 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION Sequence 994 from patent US 6605709.
ACCESSION  AR375988
VERSION    AR375988.1 GI:40079170
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 348)
AUTHORS    Breton, G.L.
TITLE      Nucleic acid and amino acid sequences relating to Proteus mirabilis
JOURNAL    for diagnostics and therapeutics
PATENT     US 6605709-A 994 12-AUG-2003;
FEATURES   Location/Qualifiers
            source
              1..348
              /organism="unknown"
              /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 11; DB 6; Length 348;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      1 TAACTATATGA 11

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 11; DB 9; Length 348;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      28 TAACTATATGA 18

RESULT 74
LOCUS      AY733638          350 bp      DNA      linear      BCT 22-SEP-2004
DEFINITION Yersinia pseudotuberculosis genomic region IG04-350.
ACCESSION  AY733638
VERSION    AY733638.1 GI:52222782
KEYWORDS   Yersinia pseudotuberculosis
SOURCE     Yersinia pseudotuberculosis
ORGANISM   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Yersinia.
REFERENCE  1 (bases 1 to 350)
AUTHORS    Wang, X. and Yang, R.
TITLE      Comparative genomics analysis of Yersinia pestis and Yersinia
JOURNAL    pseudotuberculosis using suppression subtractive hybridization and
            DNA microarray
            Unpublished
            2 (bases 1 to 350)
            Wang, X. and Yang, R.
            Direct Submission
            Submitted (25-AUG-2004) Laboratory of Analytical Microbiology,
            Institute of Microbiology and Epidemiology, Academy of Military
            Medical Sciences, No. 20, Dongdajie, Fentai, Beijing 100071, China

```

```

|||||
6 TAACTATATGA 16

```

```

RESULT 73
LOCUS      HS211WA9          348 bp      DNA      linear      PRI 28-NOV-1994
DEFINITION H. sapiens (D8S513) DNA segment containing (CA) repeat; clone
            AFM211wa9; single read.
ACCESSION  Z23603
VERSION    Z23603.1 GI:393803
KEYWORDS   CA repeat; dinucleotide repeat; GT repeat; microsatellite DNA;
            microsatellite marker; repeat polymorphism.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
AUTHORS    Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C.,
            Millasseau, P., Marc, S., Bernardi, G., Lathrop, M. and Weissenbach, J.
            The 1993-94 Genethon human genetic linkage map
            Nat. Genet. 7 (2 Spec No), 246-339 (1994)
            95004593
            7545953
            2 (bases 1 to 348)
            Weissenbach, J.
            Direct Submission
            Submitted (12-JUL-1993) Genethon, B.P. 60, 91002 Evry Cedex France.
            E-mail: Jean.Weissenbach@genethon.fr
            cloning vector is M13mp18ASBB;
            full automatic.
FEATURES   Location/Qualifiers
            source
              1..348
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="8"
              /cell_line="CEPH 134702"
              /clone_lib="genomic DNA"

```

```

FEATURES   Location/Qualifiers
            source
              1..348
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="8"
              /cell_line="CEPH 134702"
              /clone_lib="genomic DNA"

```

```

ORIGIN

```

```

Query Match      100.0%; Score 11; DB 9; Length 348;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 TAACTATATGA 11
      |||||
Db      28 TAACTATATGA 18

```

```

RESULT 74
LOCUS      AY733638          350 bp      DNA      linear      BCT 22-SEP-2004
DEFINITION Yersinia pseudotuberculosis genomic region IG04-350.
ACCESSION  AY733638
VERSION    AY733638.1 GI:52222782
KEYWORDS   Yersinia pseudotuberculosis
SOURCE     Yersinia pseudotuberculosis
ORGANISM   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Yersinia.
REFERENCE  1 (bases 1 to 350)
AUTHORS    Wang, X. and Yang, R.
TITLE      Comparative genomics analysis of Yersinia pestis and Yersinia
JOURNAL    pseudotuberculosis using suppression subtractive hybridization and
            DNA microarray
            Unpublished
            2 (bases 1 to 350)
            Wang, X. and Yang, R.
            Direct Submission
            Submitted (25-AUG-2004) Laboratory of Analytical Microbiology,
            Institute of Microbiology and Epidemiology, Academy of Military
            Medical Sciences, No. 20, Dongdajie, Fentai, Beijing 100071, China

```

```

REFERENCE  1 (bases 1 to 350)
AUTHORS    Wang, X. and Yang, R.
TITLE      Direct Submission
JOURNAL    Submitted (25-AUG-2004) Laboratory of Analytical Microbiology,
            Institute of Microbiology and Epidemiology, Academy of Military
            Medical Sciences, No. 20, Dongdajie, Fentai, Beijing 100071, China

```

```

JOURNAL    Submitted (25-AUG-2004) Laboratory of Analytical Microbiology,
            Institute of Microbiology and Epidemiology, Academy of Military
            Medical Sciences, No. 20, Dongdajie, Fentai, Beijing 100071, China

```

```

Qy      1 TAACTATATGA 11

```

```
FEATURES
  source      Location/Qualifiers
              1..350
              /organism="Yersinia pseudotuberculosis"
              /mol_type="genomic DNA"
              /strain="ATCC 29833"
              /db_xref="taxon:633"
  misc_feature 1..350
              /note="genomic region iG04-350"

ORIGIN
Query Match      100.0%; Score 11; DB 1; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
        |||||||
Db      262 TAACTATATGA 272

RESULT 75
AR395484
LOCUS      AR395484      351 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION Sequence 1499 from patent US 6617156.
ACCESSION AR395484
VERSION   AR395484.1 GI:40124117
KEYWORDS
SOURCE     Unknown.
ORGANISM   Unknown.
            Unclassified.
REFERENCE  1 (bases 1 to 351)
AUTHORS   Doucette-Stamm L.A. and Bush, D.
TITLE      Nucleic acid and amino acid sequences relating to Enterococcus
            faecalis for diagnostics and therapeutics
JOURNAL    Patent: US 6617156-A 1499 09-SEP-2003;
FEATURES   Location/Qualifiers
            1..351
            source      /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 11; DB 6; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
        |||||||
Db      156 TAACTATATGA 166
```

Search completed! August 30, 2005, 09:49:48
Job time : 55.1206 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 03:53:32 ; Search time 38.5461 Seconds
(without alignment)
10862.499 Million cell updates/sec

Title: US-09-983-000a-1_COPY_1262_1272

Perfect score: 11

Sequence: 1 taactatatga 11

Scoring table: OLIGO NUC

Gapop_60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	100.0	40	9	CG722547
2	11	100.0	43	1	AT910779
3	11	100.0	52	9	CR170566
4	11	100.0	60	8	BH414023
5	11	100.0	64	1	AI393080
6	11	100.0	66	5	CD842679
7	11	100.0	67	5	BQ763384
8	11	100.0	70	1	AA289328
9	11	100.0	73	6	CD842954
10	11	100.0	80	6	CD952670
11	11	100.0	90	6	CD400238
12	11	100.0	95	9	AJ597359
13	11	100.0	101	2	BF114252
14	11	100.0	103	8	CC029121
15	11	100.0	104	9	EX998122
16	11	100.0	105	1	AI779200
17	11	100.0	110	2	AW813864
18	11	100.0	110	8	AZ848667
19	11	100.0	110	9	CG792826
20	11	100.0	115	1	AU008489
21	11	100.0	115	8	BH919491
22	11	100.0	116	4	BJ440834
23	11	100.0	122	8	AZ894079
24	11	100.0	124	2	AW864023
25	11	100.0	125	8	BZ289448
26	11	100.0	126	2	AW868991
27	11	100.0	126	2	BE321034
28	11	100.0	128	2	AW864065
29	11	100.0	129	5	EX088472
30	11	100.0	129	7	CV390543
31	11	100.0	129	9	CR399133
32	11	100.0	130	2	AW864041
33	11	100.0	130	4	BG161052
34	11	100.0	130	8	BZ779862
35	11	100.0	131	5	BW114667
36	11	100.0	131	7	CR772977
37	11	100.0	131	9	CR399132
38	11	100.0	132	7	CK548833
39	11	100.0	132	8	AQ479469
40	11	100.0	137	5	BQ975849
41	11	100.0	137	9	EX547498
42	11	100.0	138	8	BH845619
43	11	100.0	138	9	AL950353
44	11	100.0	138	9	CL523106
45	11	100.0	141	1	AA646525
46	11	100.0	141	1	AV855814
47	11	100.0	141	4	BG370310
48	11	100.0	143	5	BQ372230
49	11	100.0	143	5	BU030907
50	11	100.0	143	8	BH685321
51	11	100.0	143	8	BH685321
52	11	100.0	147	2	AW864019
53	11	100.0	148	2	AW366932
54	11	100.0	148	5	BU034947
55	11	100.0	148	5	BU012495
56	11	100.0	149	8	BH012568
57	11	100.0	150	7	CK923867
58	11	100.0	150	8	BZ779861
59	11	100.0	151	8	BH142974
60	11	100.0	151	8	BH849360
61	11	100.0	152	6	CB356687
62	11	100.0	156	5	BQ973218
63	11	100.0	157	1	AA998828
64	11	100.0	158	8	BH142169
65	11	100.0	159	5	BQ978178
66	11	100.0	159	8	B36393
67	11	100.0	160	1	AA905814
68	11	100.0	160	9	AL765938
69	11	100.0	161	2	BB569554
70	11	100.0	161	7	CO001091
71	11	100.0	161	9	CR049920
72	11	100.0	163	9	CL927037
73	11	100.0	165	7	CO319517
74	11	100.0	165	9	BX536150
75	11	100.0	166	8	AZ112443
76	11	100.0	168	2	BE056288
77	11	100.0	168	7	CV372838
78	11	100.0	168	8	BH005149
79	11	100.0	169	5	BQ357780
80	11	100.0	169	8	BZ664557
81	11	100.0	170	1	AA933316
82	11	100.0	170	8	CC344465
83	11	100.0	170	9	CG786371
84	11	100.0	171	1	AA990242
85	11	100.0	171	7	CO908580
86	11	100.0	171	9	CR361004
87	11	100.0	172	5	BQ357776
88	11	100.0	175	1	AA501396
89	11	100.0	175	9	BX891123
90	11	100.0	177	8	AZ427505
91	11	100.0	179	8	BZ613025
92	11	100.0	179	9	AJ591228
93	11	100.0	180	7	CN621759
94	11	100.0	180	9	AJ596834
95	11	100.0	181	8	AZ359464
96	11	100.0	181	8	CC098957
97	11	100.0	181	9	CR361005
98	11	100.0	181	9	CR361005
99	11	100.0	181	9	CR361005
100	11	100.0	181	9	CR361005

c 98 11 100.0 182 7 C0911842 C0911842 BJ03024E1
 c 99 11 100.0 183 1 AA828679 AA828679 od75d12.s
 c 100 11 100.0 183 7 C0001103 C0001103 L0F38 pr

ALIGNMENTS

RESULT 1
 CG722547 1119072E05.2EL.x1 1119 - RescueMu Grid AA Zea mays genomic, genomic
 LOCUS 40 bp DNA linear GSS 20-OCT-2003
 DEFINITION survey sequence.

ACCESSION CG722547
 VERSION CG722547.1 GI:37757510
 KEYWORDS GSS.

SOURCE Zea mays
 ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
 TITLE clade; Panicoidae; Andropogoneae; Zea.
 JOURNAL 1 (bases 1 to 40)

COMMENT Walbot, V.
 Maize genomic sequences found using engineered RescueMu transposon
 Unpublished (2001)

CONTACT: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221

Email: walbot@stanford.edu

Possible ligation site of ends cut by 2 different endonucleases.
 Reverse complemented post-ligation sequence from source sequence.
 Plate: 1119072 row: 29
 Class: transposon-tagged.

FEATURES
 source Location/Qualifiers

1..40
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="mixed background W23/A188/B73/K55"
 /db_xref="taxon:4577"
 /tissue_type="leaf"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="1119 - RescueMu Grid AA"
 /note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid AA was grown at UC San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 100.0%; Score 11; DB 9; Length 40;
 Best Local Similarity 100.0%; Pred. No. 9.9e+03; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0

QY 1 TAACTATATGCA 11
 |||||
 DB 10 TAACTATATGCA 20

RESULT 2
 A1910779 4141h10.x1 NCI_CGAP_Col6 Homo sapiens cDNA clone IMAGE:2392867 3'
 LOCUS 43 bp mRNA linear EST 21-DEC-1999
 DEFINITION

similar to TR:Q29263 Q29263 UNKNOWN PROTEIN ;, mRNA sequence.
 A1910779
 VERSION A1910779.1 GI:5630515
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 1 (bases 1 to 43) Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,
 M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
 Insert length: 1366 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 1.

FEATURES
 source Location/Qualifiers

1..43
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2392867"
 /tissue_type="colon tumor, RER+"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Col6"
 /note="Organ: colon; Vector: pRTT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP Col6 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1057416-1061255, and 114584-1145351).
 Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 100.0%; Score 11; DB 1; Length 43;
 Best Local Similarity 100.0%; Pred. No. 9.8e+03; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0

QY 1 TAACTATATGCA 11
 |||||
 DB 16 TAACTATATGCA 26

RESULT 3
 CR170566
 LOCUS

DEFINITION Reverse strand read from insert in 5'HPRT insertion targeting and
 chromosome engineering clone MHPN156D06, genomic survey sequence.

ACCESSION CR170566
 VERSION CR170566.1 GI:49949415
 KEYWORDS GSS; genome survey sequence; MICR.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 1 (bases 1 to 52) Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
 Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
 Rogers,J. and Bradley,A.

```

TITLE Direct Submission
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICR
FEATURES
    source
        Location/Qualifiers
            1..52
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /clone="MHPN156d06"
                /clone_lib="MHPN"

ORIGIN
Query Match 100.0%; Score 11; DB 9; Length 52;
Best Local Similarity 100.0%; Pred. No. 9.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 8 TAACTATATGA 18

RESULT 4
BH414023 60 bp DNA linear GSS 12-DEC-2001
LOCUS 1007035H07.2EL.y1 1007 - RescueMu Grid H Zea mays genomic, genomic
survey sequence.
ACCESSION BH414023
VERSION BH414023.1 GI:17591291
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 60)
Walbot,V.
Maize genomic sequences found using engineered RescueMu transposon
Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1007035 column: 15
Class: transposon-tagged.
FEATURES
    source
        Location/Qualifiers
            1..60
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /cultivar="mixed background W23/A188/B73"
                /db_xref="taxon:4577"
                /tissue_type="leaf"
                /dev_stage="adult"
                /lab_host="DH10B"
                /clone_lib="1007 - RescueMu Grid H"
                /notes="Organ: leaf; Vector: RescueMu (engineered from
                pBluescript backbone); Site 1: BamHI; Site 2: BglII;
                RescueMu is a 4.9 kb, modified maize Mu transposon
                designed to allow plasmid rescue from total genomic DNA.
                Mu elements insert preferentially into transcription web
                units. For more information on RescueMu, go to the web
                site 'www.zmdb.iastate.edu' and follow the links for
                'RescueMu.' Grid H was grown at Berkeley in 2001. DNA
                was extracted from leaf punches, double digested using
                BamHI and BglII, and ligated to form circular plasmids.
                DH10B cells were transformed and then screened on LB
                plates with ampicillin."

ORIGIN

```

```

Query Match 100.0%; Score 11; DB 8; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.4e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 44 TAACTATATGA 54

RESULT 5
AI393080 64 bp mRNA linear EST 04-FEB-1999
LOCUS t925f11.x1 NCI CGAP CLL1 Homo sapiens cDNA clone IMAGE:2109837 3'
DEFINITION similar to TR:Q29263 Q29263 UNKNOWN PROTEIN ;, mRNA sequence.
ACCESSION AI393080
VERSION AI393080.1 GI:4222627
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 64)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/hcicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
M.D., Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
FEATURES
    source
        Location/Qualifiers
            1..64
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:2109837"
                /tissue_type="B-cell, chronic lymphocytic leukemia"
                /lab_host="DH10B"
                /clone_lib="NCI CGAP CLL1"
                /notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
                polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
                was primed with a Not I - oligo(dT) primer [5'
                TGTTACCAATCTGAAGTGGAGCGCCGATTCGTTTTTTTTTTTTTTTTTTT
                T 3']; double-stranded cDNA was ligated to Eco RI
                adaptors (Pharmacia), digested with Not I and cloned into
                the Not I and Eco RI sites of the modified pT7T3 vector.
                Library is normalized, and was constructed by Bento
                Soares and M.Fatima Bonaldo."

ORIGIN
Query Match 100.0%; Score 11; DB 1; Length 64;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 13 TAACTATATGA 23

RESULT 6
CD842679/c 66 bp mRNA linear EST 10-JUL-2003
LOCUS RFO2.128122F010914 RFO2 Brassica napus cDNA clone RFO2128122, mRNA
DEFINITION

```

```

sequence.
ACCESSION      CD842679
VERSION        CD842679.1  GI:32524619
KEYWORDS       EST.
SOURCE         Brassica napus (rape)
ORGANISM       Brassica napus
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE      1 (bases 1 to 66)
AUTHORS        Genoplante.
TITLE          Genoplante, a major partnership french program in plant genomics
JOURNAL        Unpublished (2003)
COMMENT        Contact: Genoplante
               Genoplante
               93, rue Henri Rochefort 91025 EVRY CEDEX France
               Tel: 33 1 69 47 54 00
               Fax: 33 1 69 47 54 10
               This sequence has been generated in the framework of the french
               plant genomics programme 'Genoplante' (http://www.genoplante.com
               and http://genoplante-info.infobiogen.fr).

FEATURES             source
   Location/Qualifiers
       1..66
           /organism="Brassica napus"
           /mol_type="mRNA"
           /cultivar="samourai (restored line)"
           /db_xref="taxon:3708"
           /clones="RFO2128122"
           /tissue_types="anthers"
           /clone_lib="RFO2"

ORIGIN
Query Match      100.0%; Score 11; DB 6; Length 66;
Best Local Similarity 100.0%; Pred. No. 9.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TAACTATATGA 11
Db      61 TAACTATATGA 51

RESULT 7
LOCUS      BQ763384/c
DEFINITION Ebro02 SQ008_N06 R root, 3 week, hydroponic grown, low nitrogen,
cv Optic, Ebro02 Hordeum vulgare subsp. vulgare cDNA clone
BQ763384
BQ763384.1  GI:21971856
BQ763384
EST.
SOURCE      Hordeum vulgare subsp. vulgare
ORGANISM     Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Hordeum.
REFERENCE    1 (bases 1 to 67)
AUTHORS      Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
               Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.
TITLE        Development of Barley Transcriptome Resources
JOURNAL      Unpublished (2001)
COMMENT      Contact: Waugh R, Marshall DF
               Genome Dynamics/Computational Biology
               Scottish Crop Research Institute
               Invergowrie, Dundee, DD2 5DA, Scotland, UK
               Tel: 00 44 1382 562731
               Fax: 00 44 1382 562426
               Email: est@scri.sari.ac.uk.
               Location/Qualifiers
                   1..67
                       /organism="Hordeum vulgare subsp. vulgare"
                       /mol_type="mRNA"
                       /cultivar="Optic"
                       /sub_species="vulgare"

FEATURES             source
   Location/Qualifiers
       1..70
           /organism="Mus musculus"
           /mol_type="mRNA"
           /strain="BALB/c"
           /db_xref="taxon:10090"
           /clone="IMAGE:790224"
           /sex="mixed"
           /tissue_type="kidney"
           /dev_stage="6 weeks"
           /lab_host="DH10B"
           /clone_lib="Barstead MPLRB1"
           /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
               polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA

/db_xref="taxon:112509"
/clone="EBro02_SQ008_N06"
/tissue_type="root"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="root, 3 week, hydroponic grown, low nitrogen,
cv Optic, Ebro02"
/note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from roots of 3 week old Nitrogen stressed barley
plants. Developed as part of the barley transcriptome
resources of BBSRC/SEERAD funded cereal IGF (Investigating
Gene Function) project."

ORIGIN
Query Match      100.0%; Score 11; DB 5; Length 67;
Best Local Similarity 100.0%; Pred. No. 9.2e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TAACTATATGA 11
Db      49 TAACTATATGA 39

RESULT 8
ACCESSION      AA289328/c
LOCUS          AA289328/c
DEFINITION     vc89g01.r1 Barstead MPLRB1 Mus musculus cDNA clone IMAGE:790224 5'
               similar to SW:COX3 MOUSE P00416 CYTOCHROME C OXIDASE POLYPEPTIDE
               III ; mRNA sequence.
ACCESSION      AA289328
VERSION        AA289328.1  GI:1936769
KEYWORDS       Mus musculus (house mouse)
SOURCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 70)
AUTHORS        Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
               Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
               Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
               Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
               Waterston, R.
TITLE          The WashU-HMI Mouse EST Project
JOURNAL        Unpublished (1996)
COMMENT        Contact: Marra M/Mouse EST Project
               WashU-HMI Mouse EST Project
               Washington University School
               4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
               Tel: 314 286 1800
               Fax: 314 286 1810
               Email: mouseest@watson.wustl.edu
               This clone is available royalty-free through LNL ; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               MGI:474632
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 1.
FEATURES             source
   Location/Qualifiers
       1..70
           /organism="Mus musculus"
           /mol_type="mRNA"
           /strain="BALB/c"
           /db_xref="taxon:10090"
           /clone="IMAGE:790224"
           /sex="mixed"
           /tissue_type="kidney"
           /dev_stage="6 weeks"
           /lab_host="DH10B"
           /clone_lib="Barstead MPLRB1"
           /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
               polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA

```


mRNA isolated from two-week seedlings (cultivar Kefeng 1) treated by spraying 2.0mM salicylic acid for 24, 36, 48 and 72 h. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL1-Blue MRP⁺ host cells (Stratagene)."

ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 90;

Best Local Similarity 100.0%; Pred. No. 8.9e+03; Mismatches 0; Indels 0; Gaps 0;

Matches 11; Conservative 0;

Qy 1 TAACTATATGA 11

Db 43 TAACTATATGA 33

RESULT 12

AJ597359

LOCUS

DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone 449G02, genomic survey sequence.

ACCESSION AJ597359

VERSION AJ597359.1 GI:37946987

KEYWORDS GSS; left border; T-DNA flanking sequence.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 2

AUTHORS BrunAUD, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G., Lepidiec, L., Caboche, M. and Lecharry, A.

TITLE T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites

JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)

MEDLINE 22363535

PUBMED 12446565

REFERENCE 2 (bases 1 to 95)

AUTHORS Balzerque, S.

TITLE Direct Submission

COMMENT Submitted (23-OCT-2003) Balzerque S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE

PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbsgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).

FEATURES

Location/Qualifiers

1..95

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/cultivar="Wassilewskija"

/db_xref="taxon:3702"

/clone="449G02"

/clone_lib="Arabidopsis thaliana T-DNA insertion lines"

1..95

/notes="T-DNA flanking sequence left border"

misc_feature

left border

ORIGIN

Query Match 100.0%; Score 11; DB 9; Length 95;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

Db 43 TAACTATATGA 33

RESULT 13

BF114252/c

LOCUS

DEFINITION

esculentum cDNA clone cLEY26119 5' sequence, mRNA sequence.

ACCESSION BF114252

VERSION BF114252.1 GI:10943942

KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum

REFERENCE 1 (bases 1 to 101)

AUTHORS van der Hoeven, R.S., Kochian, L., Garvin, D., Matern, A.L., Holt, I.E., Liang, F., Craven, M.B., Bowman, C.L., Ronning, C.M., Nierman, W., Frazer, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.

TITLE Generation of ESTs from tomato root tissue, pre-anthesis stage

JOURNAL Unpublished (2000)

CONTACT: CUGI

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>.

FEATURES

Location/Qualifiers

1..101

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="cLEY26119"

/clone_type="root"

/dev_stage="pre-anthesis stage"

/clone_lib="tomato root, plant at pre-anthesis"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Tissue supplied by Dave Garvin (USDA-ARS, Ithaca, NY 14850)."

ORIGIN

Query Match 100.0%; Score 11; DB 2; Length 101;

Best Local Similarity 100.0%; Pred. No. 8.8e+03; Mismatches 0; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

Db 88 TAACTATATGA 78

RESULT 14

CC029121/c

LOCUS

DEFINITION

genomic survey sequence.

ACCESSION CC029121

VERSION CC029121.1 GI:29443972

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE 1 (bases 1 to 103)

AUTHORS Walbot, V.

TITLE Maize genomic sequences found using engineered RescueMu transposon

JOURNAL Unpublished (2001)

COMMENT

Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu

Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 3591.1.108.1 column: 6
Class: transposon-tagged.

FEATURES

source

1. .103
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/Al88/B73/KS5"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="3591 - RescueMu Grid p"
/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site: 1: BamHI, Site 2: BglII; RescueMu is a 4.9 Kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid P was grown at Molokai in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 100.0%; Score 11; DB 8; Length 103;
Best Local Similarity 100.0%; Pred. No. 8.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 71 TAACTATATGA 61

RESULT 15

LOCUS BX998122 104 bp DNA linear GSS 05-JUL-2004
DEFINITION Reverse strand read from insert in 3'HPRT insertion targeting and chromosome engineering clone MHPPI75115, genomic survey sequence.

ACCESSION BX998122.1 GI:49729580
VERSION BX998122
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 104)
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
Direct Submission

TITLE Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES

source

1. .104
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPPI75115"
/clone_lib="MHPPI"

ORIGIN

Query Match 100.0%; Score 11; DB 9; Length 104;

Best Local Similarity 100.0%; Pred. No. 8.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 47 TAACTATATGA 57

RESULT 16

LOCUS AI779200 105 bp mRNA linear EST 18-MAY-2001
DEFINITION EST260079 tomato susceptible, Cornell Lycopersicon esculentum cDNA clone cLES7M14, mRNA sequence.

ACCESSION AI779200.1 GI:5277241
VERSION AI779200
KEYWORDS EST.

SOURCE

Lycopersicon esculentum (tomato)

ORGANISM

Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE

AUTHORS

1 (bases 1 to 105)
D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D., Giovannoni,J.J. and Martin,G.B.

TITLE

Generation of ESTs from *Pseudomonas* susceptible tomato

JOURNAL

COMMENT

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

FEATURES

source

1. .105
Location/Qualifiers
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="R11-13 (Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="cLES7M14"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato susceptible, Cornell"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2: XhoI; cLES - Tomato *Pseudomonas* Susceptible EST library. Directionally cloned cDNAs inserted into pBlueScript SK(-) at 5' end with EcoRI and 3' end with XhoI site"

ORIGIN

Query Match 100.0%; Score 11; DB 1; Length 105;
Best Local Similarity 100.0%; Pred. No. 8.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 72 TAACTATATGA 82

RESULT 17

LOCUS AW813864 110 bp mRNA linear EST 17-MAY-2000
DEFINITION RC3-ST0197-300300-016-h10 ST0197 Homo sapiens cDNA, mRNA sequence.

ACCESSION AW813864
VERSION AW813864.1 GI:7906858
KEYWORDS EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 110)
Dias Neto,E., Garcia Correa,R., Varjovski-Almeida,S., Briones,M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matcukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=at2-RC3-ST0197-300>)

300-016-h10kt3=2000-03-30kt4=1)

Seq primer: puc 18 forward

High quality sequence start: 23

High quality sequence stop: 109.

FEATURES

source

1. .110

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="ST0197"

/notes="Organ: stomach; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 11; DB 2; Length 110;

Best Local Similarity 100.0%; Pred. No. 8.7e+03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

|||||

Db 95 TAACTATATGA 85

RESULT 18

AZ848667

LOCUS

DEFINITION 2M0149E18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC2M0149E18 R, genomic survey sequence.

ACCESSION AZ848667

VERSION 1 GI:13031976

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 110)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,

Rilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von

Niederhausen, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0149 row: E column: 18

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 110.

FEATURES

source

1. .110

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0149E18"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/note="Vector: FWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pMD42 (GI:4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

ORIGIN

Query Match 100.0%; Score 11; DB 8; Length 110;

Best Local Similarity 100.0%; Pred. No. 8.7e+03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

|||||

Db 49 TAACTATATGA 59

RESULT 19

CG792826

LOCUS

DEFINITION CG792826 110 bp DNA linear GSS 07-NOV-2003

ZMMBBb0301118.r zMMBBb Zea mays genomic clone ZMMBBb0301118 3',

genomic survey sequence.

ACCESSION CG792826

VERSION CG792826.1 GI:38209162

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicaceae; Andropogoneae; Zea.

1 (bases 1 to 110)

Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J.

and Wing, R.

Sequencing of the maize genome

Unpublished (2003)

Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967
 Fax: 520 621 9288
 Email: <http://genome.arizona.edu>
 PCR Primers
 FORWARD: T7
 BACKWARD: M13r
 Plate: 0301 row: L column: 18
 Seq primer: M13r
 Class: BAC ends.

FEATURES

source
 Location/Qualifiers
 1..110
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBB0301L18"
 /lab_host="DH10B"
 /clone_lib="ZMMBB"
 /note="Vector: pBeloBAC11; Site 1: HindIII; Site 2:
 HindIII; Zea mays L. ssp. mays"

ORIGIN

Query Match 100.0%; Score 11; DB 9; Length 110;
 Best Local Similarity 100.0%; Pred. No. 8.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACATATATGA 11
 |||||
 Db 23 TAACATATATGA 33

RESULT 20

LOCUS AU008489 115 bp mRNA linear EST 31-JUL-1998
 DEFINITION AU008489 Schizosaccharomyces pombe late log phase cDNA
 Schizosaccharomyces pombe cDNA clone spc03483, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM Schizosaccharomyces pombe (fission yeast)
 Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 Schizosaccharomycetales; Schizosaccharomycetaceae;
 Schizosaccharomycetes.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1998)
 Contact: Mitsuoki Morimyo
 Genome Research Group
 National Institute of Radiological Sciences
 9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
 Email: morimyo@nirs.go.jp.

FEATURES

source

Location/Qualifiers
 1..115
 /organism="Schizosaccharomyces pombe"
 /mol_type="mRNA"
 /strain="972"
 /db_xref="taxon:4896"
 /clone="spc03483"
 /sex="h minus"
 /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
 /note="Vector: M13mp19; The cDNA library of
 Schizosaccharomyces pombe was prepared by cloning cDNA
 into the SmaI site of M13mp19 DNA and the direction of DNA
 sequences was not always from 5' to 3'. The cDNA data of
 Schizosaccharomyces pombe are available for searching on
 the World Wide Web. (URL, <http://www.nirs.go.jp>)"

ORIGIN

Query Match 100.0%; Score 11; DB 1; Length 115;
 Best Local Similarity 100.0%; Pred. No. 8.6e+03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TAACATATATGA 11
 |||||
 Db 80 TAACATATATGA 90

RESULT 21

LOCUS BH919491/c 115 bp DNA linear GSS 12-SEP-2002
 DEFINITION 3526_166_1_B07.2EL_y_1 3526 - RescueMu Grid K Zea mays genomic,
 genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Zea mays
 Zea mays
 Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 115)
 Maize genomic sequences found using engineered RescueMu transposon
 Unpublished (2001)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu

Possible ligation site of ends cut by 2 different endonucleases.

Reverse complemented post-ligation sequence from source sequence.

Plate: 3526_166_1 column: 9

Class: transposon-tagged.

FEATURES

source

Location/Qualifiers
 1..115
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="mixed background W23/A188/B73"
 /db_xref="taxon:4577"
 /tissue_type="leaf"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="3526 - RescueMu Grid K"
 /note="Organ: leaf; Vector: RescueMu (engineered from
 pBluescript backbone); Site 1: BamHI; Site 2: BglII;
 RescueMu is a 4.9 kb, modified maize Mu transposon
 designed to allow plasmid rescue from total genomic DNA.
 Mu elements insert preferentially into transcription
 units. For more information on RescueMu, go to the web
 site 'www.zmdb.iastate.edu' and follow the links for
 'RescueMu.' Grid K was grown at Molokai, Hawaii in winter
 2000-2001. DNA was extracted from leaf punches, double
 digested using BamHI and BglII, and ligated to form
 circular plasmids. DH10B cells were transformed and then
 screened on LB plates with ampicillin."

ORIGIN

Query Match 100.0%; Score 11; DB 8; Length 115;
 Best Local Similarity 100.0%; Pred. No. 8.6e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACATATATGA 11

|||||

Db 111 TAACATATATGA 101

RESULT 22

LOCUS

DEFINITION BJ440834 Dictyostelium discoideum cDNA library, VF Dictyostelium
 discoideum cDNA clone ddv44j20 3', mRNA sequence.

116 bp mRNA linear EST 13-MAR-2002

```

ACCESSION BJ440834
VERSION BJ440834.1 GI:19415556
SOURCE EST.
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
AUTHORS 1 (bases 1 to 116)
TITLE Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
JOURNAL Full length cDNA of Dictyostelium discoideum at the vegetative
COMMENT stage
CONTACT: Tadasu Shin-i
CONTACT For Genetic Resource Information
NATIONAL INSTITUTE OF GENETICS
1111 Yata, Mishima, Shizuoka 411-8540, Japan
TEL: 81-559-81-6856
FAX: 81-559-81-6855
EMAIL: tshini@genes.nig.ac.jp.
FEATURES
    source
        1..116
            /organism="Dictyostelium discoideum"
            /mol_type="mRNA"
            /strain="AX4"
            /db_xref="taxon:44689"
            /clone="ddv44j20"
            /sex="mat A"
            /dev_stage="Growth phase"
            /clone_lib="Dictyostelium discoideum cDNA library, VP"
ORIGIN
    Query Match 100.0%; Score 11; DB 4; Length 116;
    Best Local Similarity 100.0%; Pred. No. 8.6e+03;
    Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAACATATGCA 11
Db 27 TAACATATGCA 37
RESULT 23
AZ894079
LOCUS AZ894079
DEFINITION RPCI-24-182C3.TJ RPCI-24 Mus musculus genomic clone RPCI-24-182C2,
genomic survey sequence.
ACCESSION AZ894079
VERSION AZ894079.1 GI:13213024
KEYWORDS GSS.
SOURCE Mus Musculus (house mouse)
ORGANISM Mus Musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 122)
Tsengaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPCI-24
JOURNAL Unpublished (1999)
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end
page: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plate: 182 row: C column: 2
Seq primer: SP6
Class: BAC ends.
FEATURES
    Location/Qualifiers
        1..122
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="RPCI-24-182C2"
            /sex="Male"
            /cell_type="Spleen/Brain"
            /clone_lib="RPCI-24"
            /note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI; The
            RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
            library was cloned in the pTARBAC1 cloning vector at the
            BamHI sites using MboI partially digested male C57BL/6J
            DNA."
ORIGIN
    Query Match 100.0%; Score 11; DB 8; Length 122;
    Best Local Similarity 100.0%; Pred. No. 8.6e+03;
    Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAACTATATGCA 11
Db 25 TAACTATATGCA 35
RESULT 24
AW864023
LOCUS AW864023
DEFINITION PM2-SN0013-160300-001-a09 SN0013 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW864023
VERSION AW864023.1 GI:7998073
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 124)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jorgeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=&t2=PM2-SN0013-160
300-001-a09&t3=2000-03-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 83.
FEATURES
    source
        1..124
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /dev_stage="Adult"
            /clone_lib="SN0013"
            /note="Organ: stomach normal; Vector: puc18; Site 1: SmaI;
            Site 2: SmaI; A mini-library was made by cloning products
            derived from ORESTES PCR (U.S. Letters Patent application

```

No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 100.0%; Score 11; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 39 TAACTATATGA 49

RESULT 25

LOCUS BZ289448 125 bp DNA linear GSS 24-OCT-2002
DEFINITION SALK_022845.45.95.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_022845.45.95.x, genomic
survey sequence.

ACCESSION BZ289448
VERSION BZ289448.1 GI:24331105
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE BZ289448 125 bp DNA linear GSS 24-OCT-2002
AUTHORS SALK_022845.45.95.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_022845.45.95.x, genomic
survey sequence.

REFERENCE BZ289448 125 bp DNA linear GSS 24-OCT-2002
AUTHORS SALK_022845.45.95.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_022845.45.95.x, genomic
survey sequence.

TITLE A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)

JOURNAL Contact: Joseph R. Ecker
COMMENT Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379

FEATURES
source
1..125
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clones="SALK_022845.45.95.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 100.0%; Score 11; DB 8; Length 125;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 36 TAACTATATGA 46

RESULT 26

AW868991/c

LOCUS AW868991 126 bp mRNA linear EST 22-MAY-2000
DEFINITION MR3-SN0066-020500-003-f10 SN0066 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW868991
VERSION AW868991.1 GI:8003044
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 126)
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Bukaryota; Metazoa; Primates; Catarhini; Hominidae; Homo.
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, P., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

FEATURES
source
1..126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="SN0066"
/note="Organ: stomach normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

Query Match 100.0%; Score 11; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 11 TAACTATATGA 1

RESULT 27

BE321034/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

COMMENT

FEATURES

source

1..126

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="SN0066"

/note="Organ: stomach normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

Qy

Db

11

TAACTATATGA

11

TAACTATATGA

1

11

TAACTATATGA

1

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE

AUTHORS

1 (bases 1 to 126)
Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores,H.R., Imman,J.T., Weller,J.W., May,G.D. and Paiva,N.L.

TITLE

Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula root library

JOURNAL

Unpublished (2000)

COMMENT

On Jul 14, 2000 this sequence version replaced gi:9194811.

Contact: Paiva NL

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7317

Fax: 580 221 7380

Email: nlpaiva@noble.org

Medicago Genome Initiative accession: MGI:S:20342

Insert Length: 891 Std Error: 0.00

Plate: 033 row: H column: 01

Seq primer: TCACACAGGAACAGCTATGAC.

FEATURES

source

1..126

/organism="Medicago truncatula"

/mol_type="mRNA"

/db_xref="taxon:3880"

/clone="NF033H01RT"

/tissue_type="root"

/dev_stage="Developing root"

/note="vector: Lambda Zap; Total RNA was extracted from

non-modulated roots of plants grown in 1 mM nitrate

medium. Samples were taken at four time points

(approximately two days, one, two and six weeks post

germination) representing early seedling growth to

nitrogen limitation."

ORIGIN

Query Match 100.0%; Score 11; DB 2; Length 126;

Best Local Similarity 100.0%; Pred. NO. 8.5e+03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 TAACTATATGCA 11

|||||

98 TAACTATATGCA 88

RESULT 28

AW864065

DEFINITION PM2-SN0013-120400-003-g01 SN0013 Homo sapiens cDNA, mRNA linear EST 22-MAY-2000

AW864065

VERSION AW864065.1 GI:7998115

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 128)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsumura,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,

O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=&t2=PM2-SN0013-120

400-003-g01&t3=2000-04-12&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 9

High quality sequence stop: 127.

FEATURES

source

1..128

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="SN0013"

/note="Organ: stomach normal; Vector: puc18; Site_1: SmaI;

Site_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

ORIGIN

Query Match 100.0%; Score 11; DB 2; Length 128;

Best Local Similarity 100.0%; Pred. No. 8.5e+03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 TAACTATATGCA 11

|||||

43 TAACTATATGCA 53

RESULT 29

BX088472

LOCUS

DEFINITION

129 bp mRNA linear EST 10-MAY-2004

BX088472 tcav Oncorhynchus mykiss cDNA clone tcav0004c.j.09 3prim,

mRNA sequence.

ACCESSION BX088472

VERSION BX088472.2 GI:42611555

KEYWORDS EST.

SOURCE Oncorhynchus mykiss (rainbow trout)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei;

Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

REFERENCE

AUTHORS

TITLE

GOVERNMENT

UNPUBLISHED (2003)

On Jan 14, 2003 this sequence version replaced gi:27751885.

Contact: Guiguen Y

INRA - SCRIBE

Campus de beaulieu, RENNES cedex, 35042, France

Tel: 02.23.48.50.09

Fax: 02.23.48.50.20

Email: Yann.Guiguen@beaulieu.rennes.inra.fr

Sequence cleaned of vector, adaptor and repetitions. Contact us

at signenasupport@jouy.inra.fr to obtain the chromatogram of this

sequence.

Plate: 0004 row: j column: 9

Seq primer: M13F.

Location/Qualifiers

1..129

/organism="Oncorhynchus mykiss"

/mol_type="mRNA"

/db_xref="taxon:8022"

/clone="tcav0004c.j.09"
 /tissue type="adipose tissue, blood, brain,
 differentiating gonads, gills, interrenal, intestine,
 kidney, liver, muscle, ovary, pituitary, testis"
 /dev stage="from embryos to adults"
 /lab_host="DH10B"
 /clone_lib="tcav"

/note="vector: pT73D-pac; Rainbow trout multi-tissues -
 normalized + 1 subtraction ; Clone distribution : AGENAE
 Resource centre. Francois PIUMI,
 Francois.PIUMI@jouy.inra.fr, INRA, CEA Radiobiologie et
 Etude du genome (UREG), Domaine de Vilvert, 78352,
 Jouy-en-Josas cedex, France, +33 (0) 1.34.65.28.02, +33
 (0) 1.34.65.22.73"

ORIGIN

Query Match 100.0%; Score 11; DB 5; Length 129;
 Best Local Similarity 100.0%; Pred. No. 8.5e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 55 TAACTATATGA 65

RESULT 30
 CV390543
 LOCUS QV2-HT0577-010500-163-e11 HT0577 Homo sapiens cDNA, mRNA sequence.
 ACCESSION CV390543
 VERSION CV390543.1 GI:52740598
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 129)

REFERENCE
 AUTHORS Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

TITLE Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 JOURNAL 20202663
 MEDLINE 10737800
 PUBMED
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. http://www.ludwig.org.br.

FEATURES

source

1. .129
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HT0577"

/note="Organ: head neck; Vector: puc18; Site1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196, 716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 100.0%; Score 11; DB 7; Length 129;
 Best Local Similarity 100.0%; Pred. No. 8.5e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 77 TAACTATATGA 87

RESULT 31
 CR399133/c
 LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-815D01-025643,
 DEFINITION genomic survey sequence.

ACCESSION CR399133
 VERSION CR399133.1 GI:46939861
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1

REFERENCE
 AUTHORS Li,Y., Rosso,M.G., Strizhov,N., Viehoever,P. and Weisshaar,B.
 TITLE GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
 the identification of T-DNA insertion mutants in Arabidopsis
 thaliana

JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)
 MEDLINE 22755829
 PUBMED 12874060

REFERENCE
 AUTHORS Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
 Weisshaar,B.

TITLE An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
 flanking sequence tag-based reverse genetics

JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
 MEDLINE 23117147
 PUBMED 14756321

REFERENCE
 AUTHORS Strizhov,N., Li,Y., Rosso,M.G., Viehoever,P., Dekker,K.A. and
 Weisshaar,B.

TITLE High-throughput generation of sequence indexes from T-DNA
 mutagenized Arabidopsis thaliana lines

JOURNAL Biotechniques 35 (6), 1164-1168 (2003)
 PUBMED 14682050
 REFERENCE 4 (bases 1 to 129)

REFERENCE Rosso,M.G., Li,Y., Strizhov,N. and Weisshaar,B.
 TITLE Direct Submission

JOURNAL Submitted (01-MAY-2004) Weisshaar B., Max-Planck-Institut fuer
 Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
 This sequence has been recovered from the left border of the T-DNA.
 It indicates an insertion close to or within gene At5g1150.

DETAILS on the protocols used for generation of the sequence are
 described in References 1-3. The sequences are generated at the MPI
 for Plant Breeding Research in the context of the GABI-Kat project.
 GABI-Kat is part of the German Plant Genomics program designated
 'GABI'. Information on line availability can be found at:
 http://www.mpiz-koeln.mpg.de/GABI-Kat/.

FEATURES

source

1. .129
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 /ecotype="Col-0"

/note="PCR was performed on DNA from Arabidopsis thaliana
 plants (T1) which were transformed with the T-DNA from
 vector pAC106 (GenBank accession number: AJ537513). The
 lines contain one or more T-DNA insertions. The DNA
 fragment(s) resulting from the PCR were directly sequenced

to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

ORIGIN

Query Match 100.0%; Score 11; DB 9; Length 129;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 47 TAACTATATGA 37

RESULT 32
AW864041
LOCUS PM2-SN0013-120400-003-a09 SN0013 Homo sapiens cDNA, mRNA sequence.
DEFINITION AW864041
ACCESSION AW864041.1 GI:7998091
VERSION AW864041.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 130)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 130)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Laboratory of Cancer Genetics
MEDLINE Ludwig Institute for Cancer Research
PUBMED Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=st3-PM2-SN0013-120400-003-a09&st3=2000-04-12&st4=1)
Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 129.
Location/Qualifiers
1..130
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="SN0013"
/notes="Organ: stomach normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 11; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 1 TAACTATATGA 11

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 113 TAACTATATGA 103

ORIGIN

Query Match 100.0%; Score 11; DB 4; Length 130;
Best Local Similarity 100.0%; Pred. No.

JOURNAL
COMMENT

Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: ii35 row: b column: 01
Seq primer: -21M13UnivRev
Class: shotgun
High quality sequence stop: 142.
Location/Qualifiers
1. .130
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/db_xref="taxon:4558"
/clone="ii35b01"
/lab_host="DH5a"
/clone_lib="WGS-SbicolorF (DH5a methyl filtered)"
/note="Site 1: Xba I; Site 2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in M13mp19, .b/g reads in pUC19). The same ligation was transformed into DH5a."

FEATURES

source

RESULT 36
CN772977

LOCUS

DEFINITION

tae02b06.yl Hydra EST Darmstadt I Hydra magnipapillata cDNA 5',

131 bp mRNA linear EST 20-MAY-2004

mRNA sequence.

ACCESSION

CN772977

VERSION

CN772977.1 GI:47543611

KEYWORDS

EST.

SOURCE

Hydra magnipapillata

ORGANISM

Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroidea; Anthomedusae;

Hydridae; Hydra;

REFERENCE

1 (bases 1 to 131)

AUTHORS

Bode,H., Blumberg,B., Steele,R., Wigge,P., Gee,L., Nguyen,Q.,

Martinez,D., Kibler,D., Hampson,S., Clifton,S., Pape,D., Marra,M.,

Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B., Bowers,Y.,

Gibbons,M., Ritter,B., Bennett,J., Ronko,I., Tsagarishevili,R.,

Maguire,L., Kennedy,S., Waterston,R. and Wilson,R.

WashU Hydra EST Project

Unpublished (2002)

CONTACT: Hans Bode

WashU Hydra EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library was constructed by Corina Guder / GATC Konstanz, Germany

Library materials provided by Thomas Holstein / Molecular Cell

Biology, TUD, Darmstadt DNA sequencing by: Washington University

Genome Sequencing Center For information on obtaining a clone

please contact: Hans Bode (hbode@uci.edu)

Seq primer: -40UP.

FEATURES

Location/Qualifiers

1. .131

/organism="Hydra magnipapillata"

/mol_type="mRNA"

/strain="sf-1 mutant of strain 105"

/db_xref="taxon:6085"

/lab_host="TransforMax EC100 (Episcentre), T1 Phage

resistant electrocompetent cells"

/clone_lib="Hydra EST Darmstadt I"

/note="Vector: pBluescript II SK (+); Site_1: NotI;

Site_2: EcoRI"

ORIGIN

Query Match

Best Local Similarity

Matches

11; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

Qy

1 TAACTATATGCA 11

|||||

Db

110 TAACTATATGCA 120

RESULT 37

CN399132/c

LOCUS

DEFINITION

Arabidopsis thaliana T-DNA flanking sequence GK-81SD01-025605,

131 bp DNA linear GSS 02-MAY-2004

genomic survey sequence.

ACCESSION

CN399132

VERSION

CN399132.1 GI:46939860

KEYWORDS

GSS.

ORIGIN

Query Match

Best Local Similarity

Matches

11; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

Qy

1 TAACTATATGCA 11

|||||

Db

106 TAACTATATGCA 116

RESULT 35

BW114667

LOCUS

DEFINITION

BW114667 Nori Satoh unpublished cDNA library, tailbud embryo Ciona

intestinalis cDNA clone rcitb068f12 3', mRNA sequence.

131 bp mRNA linear EST 24-OCT-2002

ACCESSION

BW114667

VERSION

BW114667.1 GI:24361324

KEYWORDS

EST.

SOURCE

Ciona intestinalis

ORGANISM

Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;

Phlebobranchia; Cionidae; Ciona.

REFERENCE

1 (bases 1 to 131)

AUTHORS

Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.

TITLE

Expressed genes in Ciona intestinalis (2002c)

JOURNAL

Unpublished (2002)

COMMENT

Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES

Location/Qualifiers

1. .131

/organism="Ciona intestinalis"

/mol_type="mRNA"

/db_xref="taxon:7719"

/clone="rcitb068f12"

/tissue_type="whole animal"

/dev_stage="tailbud embryo"

/clone_lib="Nori Satoh unpublished cDNA library, tailbud embryo"

ORIGIN

```

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
1 Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P. and Weissshaar, B.
GABI-Kat Simplesearch: a flanking sequence tag (PST) database for
the identification of T-DNA insertion mutants in Arabidopsis
thaliana
JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE 22755829
PUBMED 12874060
REFERENCE
2 Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and
Weissshaar, B.
An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
flanking sequence tag-based reverse genetics
JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE 23117147
PUBMED 14756321
REFERENCE
3 Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and
Weissshaar, B.
High-throughput generation of sequence indexes from T-DNA
mutagenized Arabidopsis thaliana lines
JOURNAL Biotechniques 35 (6), 1164-1168 (2003)
PUBMED 14682050
REFERENCE
4 Strizhov, N., Li, Y., Rosso, M.G. and Weissshaar, B.
Direct Submission
JOURNAL Submitted (01-MAY-2004) Weissshaar B., Max-Planck-Institut fuer
Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
This sequence has been recovered from the left border of the T-DNA.
It indicates an insertion close to or within gene At5g21150.
Details on the protocols used for generations of the sequence are
described in References 1-3. The sequences are generated at the MPI
for Plant Breeding Research in the context of the GABI-Kat project.
GABI-Kat is part of the German Plant Genomics program designated
'GABI'. Information on line availability can be found at:
http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES
source
Location/Qualifiers
1..131
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="GK-815D01-025605"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Col-0"
/notes="PCR was performed on DNA from Arabidopsis thaliana
plants (T1) which were transformed with the T-DNA from
vector pAC106 (GenBank accession number: AJ537513). The
lines contain one or more T-DNA insertions. The DNA
fragment(s) resulting from the PCR were directly sequenced
to determine the genomic sequence flanking the insertion.
T-DNA derived sequences were removed."
ORIGIN
Query Match 100.0%; Score 11; DB 9; Length 131;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
|||||
Db 49 TAACTATATGCA 39

RESULT 38
CK548833/c 132 bp mRNA linear EST 05-OCT-2004
LOCUS swkz0.004410.z1 swk Bombyx mori cdna, mRNA sequence.
DEFINITION
ACCESSION CK548833

```

```

VERSION CK548833.1 GI:40933287
KEYWORDS EST.
SOURCE Bombyx mori (domestic silkworm)
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Bombycidae; Bombyx.
REFERENCE
1 (bases 1 to 132)
Xia, Q., Zhou, Z., Lu, C., Cheng, D., Dai, F., Li, B., Zhao, P., Zha, X.,
Cheng, T., Chai, C., Pan, G., Xu, J., Liu, C., Lin, Y., Qian, J., Hou, Y.,
Wu, Z., Li, G., Pan, M., Li, C., Shen, Y., Lan, X., Yuan, L., Li, T.,
Yu, H., Wang, G., Wan, Y., Zhu, Y., Yu, M., Shen, W., Wu, D., Xiang, Z.,
Yu, J., Wang, J., Li, R. Q., Shi, J. P., Li, H., Li, G. Y., Su, J. N.,
Wang, X. L., Li, G. Q., Zhang, Z. J., Wu, Q. F., Li, J., Zhang, Q. P., Wei, N.,
Xu, J. Z., Sun, H. B., Dong, L., Liu, D. Y., Zhao, S. L., Zhao, X. L.,
Meng, Q. S., Lan, F. D., Huang, X. G., Li, Y. Z., Fang, F., Li, C. F.,
Li, D. W., Sun, Y. Q., Zhang, Z. P., Yang, Z., Huang, Y. Q., Xi, Y., Qi, Q. H.,
He, D. D., Huang, H. Y., Zhang, X. W., Wang, Z. Q., Li, W. J., Cao, Y. Z.,
Wang, J., Ye, J., Ji, H., Li, S. T., Ni, P. X., Zhang, J. G., Zhang, Y.,
Zheng, H. K., Ye, C., Wang, J., Wong, G. K. S. and Yang, H. M.
A draft sequence for the genome of the domesticated silkworm
(Bombyx mori)
JOURNAL Unpublished (2004)
COMMENT Contact: Zhonghuai Xiang
Southwest Agricultural University
Chongqing Beibei
Tel: 86-23-68251123
Fax: 86-23-68251128
Email: xzh@swau.cq.cn.
FEATURES
source
Location/Qualifiers
1..132
/organism="Bombyx mori"
/mol_type="mRNA"
/strain="Dazhao (P50)"
/db_xref="taxon:7091"
/sex="mixed"
/dev_stage="Embryo (nondiapause)"
/clone_lib="swk"
/notes="Vector: pBluescript II SK(+)"
ORIGIN
Query Match 100.0%; Score 11; DB 7; Length 132;
Best Local Similarity 100.0%; Pred. No. 8.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
|||||
Db 11 TAACTATATGCA 1

RESULT 39
AQ479469/c 135 bp DNA linear GSS 23-APR-1999
LOCUS RPCI-11-211F2.TV RPCI-11 Homo sapiens genomic clone RPCI-11-211F2,
genomic survey sequence.
DEFINITION
ACCESSION AQ479469
VERSION AQ479469.1 GI:4661588
KEYWORDS GSS.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 135)
Zhao, S., Adams, M. D., Nierman, W., Malek, J., de Jong, P. and
Venter, J. C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200

```

Fax: 301 838 0208
 Email: hbe@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet cs (inforesgen.com). BAC end search page: http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
 Seq primer: T7
 Class: BAC ends.

FEATURES

source Location/Qualifiers

1..135
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="GDB:7580761"
 /db_xref="taxon:9606"
 /clone="RPCI-11-211P2"
 /sex="Male"
 /cell_type="Lymphocytes"
 /clone_lib="RPCI-11"
 /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; RPCI11 Human Male BAC Library"

ORIGIN

Query Match 100.0%; Score 11; DB 8; Length 135;
 Best Local Similarity 100.0%; Pred. No. 8.5e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

Db 87 TAACTATATGA 77

RESULT 40

BQ975849 LOCUS 137 bp mRNA linear EST 21-AUG-2002
 DEFINITION QH11L06.YG.ab1 QH ABCDI sunflower RHA801 Helianthus annuus CDNA clone QH11L06, mRNA sequence.

BQ975849 ACCESSION BQ975849.1 GI:22393372

VERSION BQ975849

KEYWORDS EST.

SOURCE Helianthus annuus (common sunflower)

ORGANISM Helianthus annuus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.

REFERENCE

AUTHORS 1 (bases 1 to 137)
 KoziK,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
 Lettuce and Sunflower ESTs from the Compositae Genome Project
 http://compgenomics.ucdavis.edu/
 Unpublished (2002)

TITLE Contact: Alexander KoziK [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Agamundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742

JOURNAL Fax: 1-(530)-752-9659
 COMMENT Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
 belongs to contig QH_CA_Contig3284, see <http://cgdb.ucdavis.edu/> for details.
 Plate: QH11 row: L column: 06.

FEATURES

source Location/Qualifiers
 1..137
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA801"
 /db_xref="taxon:4232"
 /clone="QH11L06"
 /lab_host="E.coli"

/clone_lib="QH ABCDI sunflower RHA801"
 /note="Vector: pBRCDNAsfIAB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transfections made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgdb.ucdavis.edu/>
 TAG TISSUE=germinating seeds
 TAG LIB=QH ABCDI sunflower RHA801
 TAG SEQ=TCGTGCGGG"

ORIGIN

Query Match 100.0%; Score 11; DB 5; Length 137;
 Best Local Similarity 100.0%; Pred. No. 8.4e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

Db 12 TAACTATATGA 22

RESULT 41

BX547498 LOCUS 137 bp DNA linear GSS 04-APR-2004
 DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-542D08-020616, genomic survey sequence.

BX547498 ACCESSION BX547498

VERSION BX547498.1 GI:32440307

KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE

AUTHORS 1
 Li,Y., Rosso,M.G., Strizhov,N., Viehoveer,P. and Weissshaar,B.
 GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana
 Bioinformatics 19 (11), 1441-1442 (2003)

JOURNAL MEDLINE 22755829

PUBMED 12874060

REFERENCE

AUTHORS 2
 Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and Weissshaar,B.

TITLE An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics
 Plant Mol. Biol. 53 (1-2), 247-259 (2003)

JOURNAL MEDLINE 23117147

PUBMED 14756321

REFERENCE

AUTHORS 3
 Strizhov,N., Li,Y., Rosso,M.G., Viehoveer,P., Dekker,K.A. and Weissshaar,B.

TITLE High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines

Biotecniques 35 (6), 1164-1168 (2003)

JOURNAL MEDLINE 14682050

PUBMED 4 (bases 1 to 137)

REFERENCE

AUTHORS 4
 Li,Y., Strizhov,N., Rosso,M.G. and Weissshaar,B.

TITLE Direct Submission

JOURNAL Submitted (31-MAR-2004) Weissshaar B., Max-Planck-Institut fuer

Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany

This sequence has been recovered from the left border of the T-DNA.

It indicates an insertion within the locus defined by BAC clone

K17E12. Details on the protocols used for generation of the

sequence are described in References 1-3. The sequences are

generated at the MPI for Plant Breeding Research in the context of

the GABI-Kat project. GABI-Kat is part of the German Plant Genomics

program designated 'GABI'. Information on line availability can be

found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES
source

Location/Qualifiers
1. .137
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-542D08-020616"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Col-0"
/notes="PCR was performed on DNA from Arabidopsis thaliana plants (Ti) which were transformed with the T-DNA from vector pAC161 (GenBank accession number: AJ537514). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

ORIGIN

Query Match 100.0%; Score 11; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 8.4e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
|||||
Db 71 TAACTATATGCA 81

RESULT 42

BH845619
LOCUS TC3-54114.TV TC3 Trypanosoma cruzi genomic clone TC3-54114, genomic survey sequence.
DEFINITION
ACCESSION BH845619
VERSION BH845619.1 GI:21416346
KEYWORDS GSS.
SOURCE Trypanosoma cruzi
ORGANISM Trypanosoma cruzi

REFERENCE

1 (bases 1 to 138)
Mylers, P.J., Aggarwal, G., Fazelinia, G., Mack, J., Marty, A., Munden, H., Nelson, S., Pentony, M., Rinta, J., Robertson, L., Seyler, A., Sisk, E., Stuart, K., Vogt, C., Worthey, E., El-Sayed, N.M., Ghedin, E. and Andersson, B.

Trypanosoma cruzi Cl-Brener TC3 BAC-end sequencing
Unpublished (2001)
Other GSSs: TC3-54114.TP
Contact: Peter Myler
Seattle Biomedical Research Institute
4 Nickerson Street, Seattle, WA 98109, USA
Tel: 206 284 8846
Fax: 206 284 0313
Email: mylerp@sebri.org

TITLE

COMMENT
Clones are derived from the Trypanosoma cruzi Cl-Brener BAC library TC3. For clone availability, please contact Dr. Bjorn Andersson at Uppsala University (bjorn.andersson@genpat.uu.se).

Seq primer: T7
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1. .138
/organism="Trypanosoma cruzi"
/mol_type="genomic DNA"
/strain="CL Brener"
/db_xref="taxon:5693"
/clone="TC3-54114"
/clone_lib="TC3"
/notes="Vector: pBelOBAC11; Site 1: Hin dIII; Constructed for Uppsala University by Marie-Christine Le Paslier in the laboratory of Denis Le Paslier at the Centre d'Etude du Polymorphisme Humain (CEPH), Paris, France. Briefly, Trypanosoma cruzi Cl-Brener agarose embedded DNA (obtained from Dr. Franco da Silveira) was partially digested with Hin dIII. High molecular weight fragments were ligated in

ORIGIN

Query Match 100.0%; Score 11; DB 8; Length 138;
Best Local Similarity 100.0%; Pred. No. 8.4e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
|||||
Db 32 TAACTATATGCA 42

RESULT 43

AL950353
LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-327H02-015979, genomic survey sequence.
DEFINITION
ACCESSION AL950353
VERSION AL950353.1 GI:24406975
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

1
Li, Y., Rosso, M.G., Strizhov, N., Viehovever, P. and Weishaar, B.
GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana
Bioinformatics 19 (11), 1441-1442 (2003)

JOURNAL

MEDLINE 22755829

PUBMED 12874060

REFERENCE

2
Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weishaar, B.

An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics
Plant Mol. Biol. 53 (1-2), 247-259 (2003)

JOURNAL

MEDLINE 23117147

PUBMED 14756321

REFERENCE

3
Strizhov, N., Li, Y., Rosso, M.G., Viehovever, P., Dekker, K.A. and Weishaar, B.

High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines
BioTechniques 35 (6), 1164-1168 (2003)

JOURNAL

PUBMED 14682050

REFERENCE

4 (bases 1 to 138)

JOURNAL

PUBMED

REFERENCE

TITLE

JOURNAL

PUBMED

COMMENT

This sequence has been recovered from the left border of the T-DNA. It indicates an insertion within the locus defined by BAC clone T19N18. Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES

source

Location/Qualifiers
1. .138
/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="GK-327H02-015979"

/clone_lib="Arabidopsis thaliana T-DNA insertion lines"

/ecotype="Col-0"

/note="PCR was performed on DNA from Arabidopsis thaliana

pBelOBAC11 digested with Hin dIII. The average insert size is 100 kb. Total clone coverage: approx. 33 X the haploid genome."


```

AV855814/c
LOCUS       AV855814               141 bp    mRNA          linear      EST 08-NOV-2001
DEFINITION   AV855814 Nori Satoh unpublished cDNA library, egg Ciona
              intestinalis cDNA clone rcieg19n13 3', mRNA sequence.
ACCESSION   AV855814
VERSION     AV855814.1   GI:16843338
KEYWORDS    EST.
SOURCE      Ciona intestinalis
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1   (bases 1 to 141)
AUTHORS     Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
TITLE       Expressed genes in Ciona intestinalis
JOURNAL     Unpublished (2000)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES             source
     source          1..141
                     /organism="Ciona intestinalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:7719"
                     /clone="rcieg19n13"
                     /tissue_type="whole animal"
                     /dev_stage="egg"
                     /clone_lib="Nori Satoh unpublished cDNA library, egg"

ORIGIN
Query Match      100.0%;   Score 11;   DB 1;   Length 141;
Best Local Similarity 100.0%;   Pred. NO. 8.4e+03;
Matches 11;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy  1  TAACTATATGCA 11
    |||||
Db  40 TAACTATATGCA 30

RESULT 47
LOCUS       BG370310               141 bp    mRNA          linear      EST 09-MAR-2001
DEFINITION   BG370310 x1 NCI_CGAP_HN20 Homo sapiens cDNA clone IMAGE:4261619 3',
              mRNA sequence.
ACCESSION   BG370310
VERSION     BG370310.1   GI:13266847
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1   (bases 1 to 141)
AUTHORS     NCI/NIH-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute / National Institute of Dental Research,
            Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
            Unpublished (1997)
JOURNAL     Other ESTs: nai30c06.Y1
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapsb@mail.nih.gov
            cDNA Library Preparation: David B. Krizman, Ph.D.
            DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
            Clone distribution by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL, send email to:
            infoimage.lnl.gov
            Seq primer: -40UP from Gibco.

FEATURES             source
     source          1..141
                     /organism="Homo sapiens"
                     /mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:4261619"
/lab_host="DH10B"
/clone_lib="NCI CGAP HN20"
/note="Organ: normal head/neck tissue; Vector: pAMP1, mRNA
made from head/neck tissue, cDNA made by oligo-dT
priming. Directionally cloned into UDG sites.
Size-selected on agarose gel, average insert size 300 bp.
Primary library. cDNA Library Preparation: David B.
Krizman, Ph.D."

ORIGIN
Query Match      100.0%;   Score 11;   DB 4;   Length 141;
Best Local Similarity 100.0%;   Pred. NO. 8.4e+03;
Matches 11;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy  1  TAACTATATGCA 11
    |||||
Db  18 TAACTATATGCA 28

RESULT 48
LOCUS       BQ372230               143 bp    mRNA          linear      EST 21-MAY-2002
DEFINITION   RC3-FN0143-190700-022-g07 FN0143 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ372230
VERSION     BQ372230.1   GI:21047744
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1   (bases 1 to 143)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,P.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL     20202663
MEDLINE     10737800
PUBMED
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&t2=RC3-FN0143-
            190700-022-g07&t3=2000-07-19&t4=1).
            Seq primer: puc 18 forward
            High quality sequence start: 13
            High quality sequence stop: 143.

FEATURES             source
     source          1..143
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="FN0143"
            /note="Organ: prostate normal; Vector: puc18; Site 1:
            SmaI; Site 2: SmaI; A mini-library was made by cloning
            products derived from ORESTES PCR (U.S. Letters Patent
            application No. 196,716 - Ludwig Institute for Cancer
            Research) profiles into the pUC 18 vector. Reverse
            transcription of tissue mRNA and cDNA amplification were
            performed under low stringency conditions."

```


ORIGIN

Query Match 100.0%; Score 11; DB 5; Length 143;
 Best Local Similarity 100.0%; Pred. No. 8.4e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 54 TAACTATATGA 64

RESULT 49

BU030907
 LOCUS BU030907 143 bp mRNA linear EST 23-AUG-2002
 DEFINITION OHJ16N09_Yg.ab1 OH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
 clone OHJ16N09, mRNA sequence.

ACCESSION BU030907.1 GI:22466427

VERSION BU030907

KEYWORDS EST.

SOURCE Helianthus annuus (common sunflower)

ORGANISM

Helianthus annuus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; campanulids; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Helianthus.

REFERENCE

1 (bases 1 to 143)
 Kozik A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
 Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
 Ellison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
 Lai, Z., Church, S., Jackson, L. and Bradford, K.

AUTHORS

Lettuce and Sunflower ESTs from the Compositae Genome Project
<http://compgenomics.ucdavis.edu/>
 Unpublished (2002)

JOURNAL

COMMENT

Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Armundson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]

belongs to contig OH_Ca_Contig3284, see <http://cgpdb.ucdavis.edu/>
 for details.

Plate: OHJ16 row: N column: 09.

FEATURES

source

1..143

/organism="Helianthus annuus"

/mol_type="mRNA"

/cultivar="RHA280"

/db_xref="taxon:4232"

/clone="OHJ16N09"

/lab_host="E.coli"

/clone_lib="OH_EFGHJ sunflower RHA280"

/note="vector: pRCNAsf1AB; The library was constructed
 from 11 different sources of RNA from a single genotype.

Separate cDNAs were generated using primers that
 incorporated unique 5' and 3' tags to distinguish each

source of RNA. cDNAs were then pooled, size-fractionated,
 directionally cloned into a custom medium-copy vector and

transformations made with four size classes to minimize
 size bias. Details of each source of RNA and library

construction can be obtained at <http://cgpdb.ucdavis.edu/>
 TAG TISSUE=germinating seeds

TAG_LIB=OH_EFGHJ sunflower RHA280

TAG_SEQ=TCGTGCGGG

ORIGIN

Query Match 100.0%; Score 11; DB 5; Length 143;
 Best Local Similarity 100.0%; Pred. No. 8.4e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 12 TAACTATATGA 22

RESULT 50

BH685321/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM4&t2=CM4-MT0023-
230600-216-g12&t3=2000-06-23&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 4
High quality sequence stop: 144.
Location/Qualifiers
source
1. 144
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="WT0023"
/note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

FEATURES
source
1. 144
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="WT0023"
/note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

ORIGIN
Query Match 100.0%; Score 11; DB 5; Length 144;
Best Local Similarity 100.0%; Pred. No. 8.4e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAACTATATGCA 11
|||||
Db 30 TAACTATATGCA 40

RESULT 52
AW864019
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
COMMENT

AW864019 147 bp mRNA linear EST 22-MAY-2000
PM2-SN0013-160300-001-c07 SN0013 Homo sapiens cDNA, mRNA sequence.
AW864019
EST.
GI:7998069
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 147)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hate,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL

FEATURES
source
1. 148
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0156"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under

FEATURES
source
1. 147
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="SN0013"
/note="Organ: stomach normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 100.0%; Score 11; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 8.4e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAACTATATGCA 11
|||||
Db 50 TAACTATATGCA 60

RESULT 53
AW366932
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
COMMENT

AW366932 148 bp mRNA linear EST 04-FEB-2000
IL0-HT0156-251099-132-c03 HT0156 Homo sapiens cDNA, mRNA sequence.
AW366932
EST.
GI:6871582
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 148)
HCGP http://www.ludwig.org.br/ORESTES.
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=IL0&t2=IL0-HT0156-
251099-132-c03&t3=1999-10-25&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 3
High quality sequence stop: 148.
Location/Qualifiers
source
1. 148
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0156"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under


```

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TAACTATATGA 11
    |||||
Db 50 TAACTATATGA 40

RESULT 59
BH142974
LOCUS
DEFINITION
TDCDU75TH cTOG Lycopersicon esculentum genomic clone CTG025N5,
genomic survey sequence.
ACCESSION
BH142974
VERSION
BH142974.1 GI:15196163
KEYWORDS
GSS.
SOURCE
Lycopersicon esculentum (tomato)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 (bases 1 to 151)
van der Hoeven,R., Sun,H., Cho,J., Utterback,T., Ronning,C. and
Tanksley,S.
Tomato Demethylated Genomic DNA Sequences
JOURNAL
Unpublished (2001)
COMMENT
Contact: CUGI
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
tomato demethylated genomic DNA
Insert Length: 1270 Std Error: 0.00
Seq primer: M13P-R
Class: shotgun.
FEATURES
source
location/Qualifiers
1..151
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/cultivar="E6203"
/db_xref="taxon:4081"
/clone="CTG025N5"
/tissue_type="young leaves"
/dev_stage="12-14 weeks post harvest"
/lab_host="E.coli JM109"
/clone_lib="cTOG"
/note="Vector: pluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; This library was made from short EcoRI digested
fragments of the genome of Lycopersicon esculentum ligated
into pBS (SK-). The fragments were cloned into the
methylation restrictive E.coli strain JM109 with the
purpose of enriching the library for non-methylated DNA
fragments. This procedure may enrich the pool of cloned
fragments in JM109 cells for sequences representing
expressed genes. Average insert size 1.27 kb."

ORIGIN
Query Match 100.0%; Score 11; DB 8; Length 151;
Best Local Similarity 100.0%; Pred. No. 8.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TAACTATATGA 11
    |||||
Db 7 TAACTATATGA 17

RESULT 60
BH849360
LOCUS
DEFINITION
SALK_069585.54.05.x Arabidopsis thaliana TDNA insertion lines
survey sequence.
ACCESSION
BH849360
VERSION
BH849360.1 GI:21420231

```

```

GSS.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 151)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jecke,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
COMMENT
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.
FEATURES
source
location/Qualifiers
1..151
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecot_type="Col-0"
/db_xref="taxon:3702"
/clone="SALK_069585.54.05.x"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN
Query Match 100.0%; Score 11; DB 8; Length 151;
Best Local Similarity 100.0%; Pred. No. 8.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TAACTATATGA 11
    |||||
Db 44 TAACTATATGA 54

RESULT 61
CB356687/c
LOCUS
DEFINITION
ZP001-P00017-DPE-P-C_H02 GISZF001 Danio rerio cDNA clone
IMAGE:6898132 5', similar to zeh10840 Zebrafish Embryonic Heart cDNA
Library Danio rerio cDNA 5', mRNA sequence.
ACCESSION
CB356687
VERSION
CB356687.1 GI:28997555
KEYWORDS
EST.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 152)
Mathavan,S., Wei,C., Thoreau,H., Chia,J.M. and Ruan,Y.
Genome Institute of Singapore, Zebrafish EST Collection
Unpublished (2003)
COMMENT
Contact: Ruan Y
Cloning and Sequencing
Genome Institute of Singapore
60 Biopolis Street, #02-01, Genome, Singapore 138672
Tel: +65 6478 8073
Fax: +65 6478 9059
Email: ruanyj@gis.a-star.edu.sg

```

GIS Clone ID: ZF001-P00017-PP_P03

PCR Primers

FORWARD: M13

BACKWARD: M13

Plate: ZF001-P00017-DPE-F-C

Seq primer: CCGCATAACTGTATAGCA

High quality sequence stop: 152.

FEATURES

Location/Qualifiers

1. .152

/organism="Danio rerio"

/mol_type="mRNA"

/db_xref="taxon:7955"

/clone="IMAGE:6898132"

/tissue_type="Embryo"

/dev_stage="7 Different embryonic Stages (From just

fertilized Embryos to 72 hours just hatched baby fish)"

/lab_host="DH10B"

/clone_lib="GISZF001"

/notes="Vector: pDNR-LIB; Site 1: Sfi A (GGCATTACGGCC);

Site 2: Sfi B (GGCCCTCGGCC); Priming method: Sfi-(dT)30

Primed ; Priming sequence: 5.ATTCTAGA GGCAGGCGGCC

GACATG(T)30VN ; Directionally cloned, 5' linker/adaptor

sequence: Sfi A site GGCATTACGGCC ; 3' cloning site: Sfi B

site 5.AAGCAGTCGATCAACGACAGATGGCC ; 3' linker/adaptor

sequence: same

as the priming sequence ; Average insert size: 2kb ; For

PCR insert analysis: Use M13 Forward and reverse primers ;

Library Amplified Recombinants (inserts): 98 ; Library

complexity: 5x10⁶ ; Full-length construction (method):

SMART, a Clontech method ; Library constructed by: S.

Mathavan, Chia-Lin Wei, and Yijun Ruan Genome Institute of

Singapore"

ORIGIN

Query Match 100.0%; Score 11; DB 6; Length 152;

Best Local Similarity 100.0%; Pred. No. 8.3e+03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

|||||

Db 113 TAACTATATGA 103

RESULT 62

BQ973218

LOCUS

DEFINITION BQ973218.1 yg.ab1 QH ABCDI sunflower RHA801 Helianthus annuus cDNA

clone QH112D19, mRNA sequence.

ACCESSION BQ973218

VERSION BQ973218.1 GI:22390741

KEYWORDS EST.

SOURCE Helianthus annuus (common sunflower)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; campanulids; Asterales; Asteraceae; Asteroideae;

Heliantheae; Helianthus.

1 (bases 1 to 156)

Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,

Lin,H., van Damme,M., Lavelle,D., Chevallier,P., Ziegler,J.,

Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,

Lai,Z., Church,S., Jackson,L. and Bradford K.

Lettuce and Sunflower ESTs from the Compositae Genome Project

http://compenomics.ucdavis.edu/

Unpublished (2002)

Contact: Alexander Kozik [R.W.Michelmore]

Department of Vegetable Crops, R.W.Michelmore Lab

University of California at Davis (UCD)

Amundson Hall, UCD, Davis, CA 95616, USA

Ref: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu]

belongs to contig QH_CA_Contig3284, see http://cgdb.ucdavis.edu/

for details

Plate: QH112 row: D column: 19.

FEATURES

source

1. .156

/organism="Helianthus annuus"

/mol_type="mRNA"

/cultivar="RHA801"

/db_xref="taxon:4232"

/clone="QH112D19"

/lab_host="E.coli"

/clone_lib="QH ABCDI sunflower RHA801"

/note="Vector: pBRCDNASfiAB; The library was constructed

from 11 different sources of RNA from a single genotype.

Separate cDNAs were generated using primers that

incorporated unique 5' and 3' tags to distinguish each

source of RNA. cDNAs were then pooled, size-fractionated,

directionally cloned into a custom medium-copy vector and

transformations made with four size classes to minimize

size bias. Details of each source of RNA and library

construction can be obtained at http://cgdb.ucdavis.edu/

TAG_SEQ=Not found"

ORIGIN

Query Match 100.0%; Score 11; DB 5; Length 156;

Best Local Similarity 100.0%; Pred. No. 8.3e+03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

|||||

Db 31 TAACTATATGA 41

RESULT 63

AA999828

LOCUS

DEFINITION

os42f09.ab1 NCI CGAP Br2 Homo sapiens cDNA clone IMAGE:1608041 3'

similar to TR:Q29263 Q29263 UNKNOWN PROTEIN ; mRNA sequence.

ACCESSION AA999828

VERSION AA999828.1 GI:3190383

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 157)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

cDNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LINL at:

www-bio.lnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1. .157

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1608041"

/sex="female, pooled"

/tissue_type="breast"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_Br2"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled bulk breast tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. This library is the normalized version of NCI CGAP Br1.1. Library was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 100.0%; Score 11; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 8.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 16 TAACTATATGA 26

RESULT 64
BH142169 158 bp DNA linear GSS 16-AUG-2001
LOCUS
DEFINITION TDG6181TH cTOG Lycopersicon esculentum genomic clone cTOG23M18,
genomic survey sequence.

ACCESSION BH142169
VERSION BH142169.1 GI:15194451
KEYWORDS GSS.
SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 158)
REFERENCE van der Hoeven, R., Sun, H., Cho, J., Utterback, T., Ronning, C. and Tanksley, S.
AUTHORS

Tomato Demethylated Genomic DNA Sequences
Unpublished (2001)
Contact: CUGI

COMMENT
Clemson University
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
tomato demethylated genomic DNA
Insert Length: 1270 Std Error: 0.00
Seq primer: M13P-R
Class: shotgun.

FEATURES

source
1..158
Location/Qualifiers
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/cultivar="B6203"
/db_xref="taxon:4081"
/clone="cTOG23M18"
/tissue_type="young leaves"
/dev_stage="12-14 weeks post harvest"
/lab_host="E.coli JM109"
/clone_lib="cTOG"

/note="Vector: pluescript SK(-); Site 1: EcoRI; Site 2: XhoI; This library was made from short EcoRI digested fragments of the genome of Lycopersicon esculentum ligated into pBS (SK-). The fragments were cloned into the methylation restrictive E.coli strain JM109 with the purpose of enriching the library for non-methylated DNA fragments. This procedure may enrich the pool of cloned fragments in JM109 cells for sequences representing expressed genes. Average insert size 1.27 kb."

ORIGIN

Query Match 100.0%; Score 11; DB 8; Length 158;
Best Local Similarity 100.0%; Pred. No. 8.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 7 TAACTATATGA 17

RESULT 65
BQ978178/c

LOCUS
DEFINITION BQ978178 159 bp mRNA linear EST 21-AUG-2002
clone QH14C12, mRNA sequence.

ACCESSION BQ978178
VERSION BQ978178.1 GI:22395701
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)

ORGANISM Helianthus annuus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.

REFERENCE 1 (bases 1 to 159)

AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkmann, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
<http://compgenomics.ucdavis.edu/>
Unpublished (2002)

JOURNAL
COMMENT

Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Aamundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659

Email: akozikeatgc.org [michelmore@vegmail.ucdavis.edu]
Belongs to contig QH_CA_Contig3554, see <http://cgdb.ucdavis.edu/>
for details.
Plate: QH14 row: C column: 12..

FEATURES

source
1..159
Location/Qualifiers
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA801"
/db_xref="taxon:4232"
/clone="QH14C12"
/lab_host="E.coli"
/clone_lib="QH-ABCDI sunflower RHA801"
/note="Vector: pBRCNDASFIAB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgdb.ucdavis.edu/>
TAG_SEQ=Not found"

ORIGIN

Query Match 100.0%; Score 11; DB 5; Length 159;
Best Local Similarity 100.0%; Pred. No. 8.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 153 TAACTATATGA 143

RESULT 66
B36393/c

LOCUS
DEFINITION B36393 159 bp DNA linear GSS 17-OCT-1997
HS-1040-B2-A01-MR.abi CIT Human Genomic Sperm Library C Homo
sapiens genomic clone Plate=CT 822 Col=2 Row=B, genomic survey
sequence.

```

ACCESSION B36393
VERSION B36393.1 GI:2535762
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 159)
AUTHORS Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
Traicoiff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
TITLE Construction of a Characterized Clone Resource for Genomic
Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
Tagged Connectors
JOURNAL Unpublished (1997)
COMMENT Contact: Mahairas GG, Zackrone KD, Hood L
University of Washington
Seattle, WA 98195, USA
Tel: (206) 616-8744
Fax: (206) 685-7301
Email: kzackron@u.washington.edu
Sequence Tagged Connector
Plate#: CT 822 row: B column: 2
Class: BAC ends
High quality sequence stop: 159.
FEATURES             source
    source            1..159
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /clones="Plate=CT 822 Col=2 Row=B"
                        /sex="M"
                        /clone_lib="CIT Human Genomic Sperm Library C"
                        /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
                        E-Coli DH10B"

ORIGIN
Query Match          100.0%; Score 11; DB 8; Length 159;
Best Local Similarity 100.0%; Pred. No. 8.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACATATATGA 11
    |||||
Db 51 TAACATATATGA 41

RESULT 67
LOCUS AA905814
DEFINITION AA904814 160 bp mRNA linear EST 09-JUN-1998
INAG#:1505131 3' similar to TR:Q29263 Q29263 UNKNOWN PROTEIN ; ,
mRNA sequence.
ACCESSION AA904814
VERSION AA904814.1 GI:3040937
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 160)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumot Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 817 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 1.
FEATURES             source
    source            1..160
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone_lib="DH10B"
                        /note="Organ: pooled; Vector: p77T3D-Pac (Pharmacia) with
                        a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
                        Equal amounts of plasmid DNA from three normalized
                        libraries (fetal lung NbH19W, testis NHT, and B-cell
                        NCI-CCAP GCBI) were mixed, and as circles were made in
                        vitro. Following HAP purification, this DNA was used as
                        tracer in a subtractive hybridization reaction. The driver
                        was PCR-amplified cDNAs from pools of 5,000 clones made
                        from the same 3 libraries. The pools consisted of
                        I.M.A.G.E. clones 297480-302087, 682632-687239,
                        726408-728711, and 729096-731399. Subtraction by Bento
                        Soares and M. Fatima Bonaldo. "

ORIGIN
Query Match          100.0%; Score 11; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 8.3e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACATATATGA 11
    |||||
Db 16 TAACATATATGA 26

RESULT 68
LOCUS AL765938/c
DEFINITION AL765938 160 bp DNA linear GSS 01-APR-2004
Arabidopsis thaliana T-DNA flanking sequence GK-145C09-012979,
genomic survey sequence.
ACCESSION AL765938
VERSION AL765938.1 GI:21519071
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Li,Y., Rosso,M.G., Strizhov,N., Viehoveer,P. and Weishaar,B.
TITLE GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
the identification of T-DNA insertion mutants in Arabidopsis
thaliana
JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE 22755829
PUBMED 12874060
REFERENCE 2
AUTHORS Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
Weishaar,B.
TITLE An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
flanking sequence tag-based reverse genetics
JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE 23117147
PUBMED 14756321
REFERENCE 3
AUTHORS Strizhov,N., Li,Y., Rosso,M.G., Viehoveer,P., Dekker,K.A. and
Weishaar,B.
TITLE High-throughput generation of sequence indexes from T-DNA
mutagenized Arabidopsis thaliana lines
JOURNAL Biotechniques 35 (6), 1164-1168 (2003)
PUBMED 14682050
REFERENCE 4 (bases 1 to 160)
AUTHORS Rosso,M.G., Li,Y., Strizhov,N. and Weishaar,B.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2004) Weishaar B., Max-Planck-Institut fuer
Zuechtungsforshung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT This sequence has been recovered from the left border of the T-DNA.
It indicates an insertion close to or within gene At4g25070.
Details on the protocols used for generation of the sequence are

```


Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. *Genome Res.* 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303,
19-44 (1999)

```

Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.
Location/Qualifiers
1..161
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="3830403D08"
/sex="female"
/tissue_type="placenta and extra embryonic tissue"
/dev_stages="18 days pregnant, adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 18 days pregnant,
placenta and extra embryonic tissue"

```

prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAAGGATCCCAAGAGCTCTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse

cDNA was prepared with the primer adapter of sequence 5'-GAGAGAGAGATTCGAGTCTAATTAATTAATTCCTCCCCCCC 3'. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified phluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

```

1  TAACTATATGGA 11
   |||||
   100.0%; Score 11; DB 2; Length 161;
   1 Similarity 100.0%; Pred. No. 8.3e+03;
   11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

5

CO001091 161 bp mRNA linear EST 09-JUN-2004
 LO25 pre-oestrus adult sheep ovary library Ovis aries cDNA, mRNA
 sequence.
 CO001091
 CO001091.1 GI:48507980
 EST.

KEYWORDS

Ovis aries (sheep)
Ovis aries
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Caprinae; Ovis.
1 (bases 1 to 151)
Liu, S.-F. and Du, L.-X.
Gene expression in sheep ovary
Unpublished (2004)
Contact: LIU, S.-F., DU, L.-X.
Laboratory of Animal Biotechnology
Shandong Agricultural University
61 Daizong Street, Tainan, Shandong Province 271018, China;
Institute of Animal Husbandry, Chinese Academy of Agricultural
Science, Beijing 100094, China
Tel: 01062819997

```

Email: lxdu0263.net.
FEATURES
  source
    Location/Qualifiers
      1..161
        /organism="Ovis aries"
        /mol_type="mRNA"
        /strain="Little-tailed Han Sheep"
        /db_xref="taxon:9940"
        /sex="female"
        /dev_stage="pre-oestrus"
        /lab_host="BM25.8"
        /clone_lib="pre-oestrus adult sheep ovary library"
        /notes="Organ: ovary; Vector: TriplEx2; Site 1: Sfi I;
        Site 2: Sfi I; Non-normalized library, sequenced 3' with
        TriplEx2 primer(TAATACCACTACATAGGG). Library constructed
        with total RNA extracted using the Trizol method and
        pooled from 3 females."
ORIGIN
  Query Match      100.0%; Score 11; DB 7; Length 161;
  Best Local Similarity 100.0%; Pred. No. 8.3e+03;
  Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAACTATATGCA 11
    |||||
Db 60 TAACTATATGCA 50

RESULT 71
CR049920/c
LOCUS
DEFINITION
  Reverse strand read from insert in 5'HPRT insertion targeting and
  chromosome engineering clone MHPN154115, genomic survey sequence.
ACCESSION
  CR049920
VERSION
  CR049920.1 GI:49783059
KEYWORDS
  GSS; genome survey sequence; MICER.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 161)
  Adams,D.J., Biggs,F.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
  Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
  Rogers,J. and Bradley,A.
  Direct Submission
  Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
  CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES
  source
    Location/Qualifiers
      1..161
        /organism="Mus musculus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10090"
        /clone_lib="MHPN154115"
        /clone_lib="MHPN"
ORIGIN
  Query Match      100.0%; Score 11; DB 9; Length 161;
  Best Local Similarity 100.0%; Pred. No. 8.3e+03;
  Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAACTATATGCA 11
    |||||
Db 129 TAACTATATGCA 119

RESULT 72
CL927037
LOCUS
DEFINITION
  OA_Aba0033M13.f OA_Aba Oryza australiensis genomic clone
  OA_Aba0033M13 5', genomic survey sequence.
ACCESSION
  CL927037
VERSION
  CL927037.1 GI:52048297
KEYWORDS
  GSS.
SOURCE
  Oryza australiensis
    ORGANISM
      Oryza australiensis
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
      Ehrhartoideae; Oryzeae; Oryza.
      REFERENCE
        1 (bases 1 to 163)
        Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
        Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
        OMAP Project
        Unpublished (2004)
        JOURNAL
        COMMENT
          Contact: Rod A. Wing
          Arizona Genomics Institute
          University of Arizona
          Forbes Building Room 303, Tucson, AZ 85721-0036, USA
          Tel: 520 626 9595
          Fax: 520 621 1259
          Email: http://genome.arizona.edu
          PCR Primers
          FORWARD: TAA TAC GAC TCA CTA TAG GG
          BACKWARD: CAC TCA TTA GGC ACC CCA
          Plate: 0033 row: M column: 13
          Seq primer: TAA TAC GAC TCA CTA TAG GG
          Class: BAC ends.
          Location/Qualifiers
            1..163
              /organism="Oryza australiensis"
              /mol_type="genomic DNA"
              /db_xref="taxon:4532"
              /clone_lib="OA_Aba0033M13"
              /tissue_type="young leaves"
              /lab_host="DH10B T1 phage resistant"
              /clone_lib="OA_Aba"
              /note="Vector: pGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN
  Query Match      100.0%; Score 11; DB 9; Length 163;
  Best Local Similarity 100.0%; Pred. No. 8.3e+03;
  Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAACTATATGCA 11
    |||||
Db 99 TAACTATATGCA 109

RESULT 73
CO319517
LOCUS
DEFINITION
  EK278444.5prime Exelixis FlyTag CK01 pCDNA-SK+ Drosophila
  melanogaster cDNA clone EK278444 5, mRNA sequence.
ACCESSION
  CO319517
VERSION
  CO319517.1 GI:49377951
KEYWORDS
  EST.
SOURCE
  Drosophila melanogaster (fruit fly)
ORGANISM
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
  REFERENCE
    1 (bases 1 to 165)
    Kopczyński,C., Platt,D., Campbell,J., Muzong,C., Laufer,A.,
    Peterson,E. and Swimmer,C.
    Exelixis FlyTag EST Project CK01 Library
    Unpublished (2004)
    JOURNAL
    COMMENT
      Contact: Stapleton, M.
      BDGP
      Lawrence Berkeley National Lab
      One Cyclotron Rd, Berkeley, CA 94720, USA
      Fax: 510 486 6798
      Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
      Plate: EK 2784 row: D column: 8
      High quality sequence stop: 151.
      Location/Qualifiers
        1..165
          /organism="Drosophila melanogaster"
          /mol_type="mRNA"
FEATURES
  source
    Location/Qualifiers
      1..165
        /organism="Drosophila melanogaster"
        /mol_type="mRNA"

```

```

/db_xref="taxon:7227"
/clone="EK278444"
/clone_lib="Exelixis FlyTag CK01 pCDNA-SK+"
/notes="Organ: mixed stage embryos; imaginal disks, and
adult heads; Vector: pCDNA-SK+; Site 1: NotI; Site 2:
XhoI; Random primed, normalized library from mixed stage
embryos, imaginal disks, and adult heads."

ORIGIN
Query Match      100.0%; Score 11; DB 7; Length 165;
Best Local Similarity 100.0%; Pred. No. 8.2e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGCA 11
        |||||
Db      122 TAACTATATGCA 132

RESULT 74
BX536150
LOCUS      BX536150
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-523H09-020210,
            genomic survey sequence.
ACCESSION  BX536150
VERSION     BX536150.1 GI:31413280
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE   1
AUTHORS     Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P. and Weisshaar, B.
TITLE       GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
            the identification of T-DNA insertion mutants in Arabidopsis
            thaliana
JOURNAL     Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE     22755829
PUBMED      12874060

REFERENCE   2
AUTHORS     Rosso, M.G., Li, Y., Strizhov, N., Reles, B., Dekker, K. and
            Weisshaar, B.
TITLE       An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
            flanking sequence tag-based reverse genetics
JOURNAL     Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE     23117147
PUBMED      14756321

REFERENCE   3
AUTHORS     Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and
            Weisshaar, B.
TITLE       High-throughput generation of sequence indexes from T-DNA
            mutagenized Arabidopsis thaliana lines
JOURNAL     Biotechniques 35 (6), 1164-1168 (2003)
MEDLINE     14682050
PUBMED      14682050

REFERENCE   4
AUTHORS     Rosso, M.G., Li, Y., Strizhov, N. and Weisshaar, B.
TITLE       Direct Submission
JOURNAL     Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
            This sequence has been recovered from the left border of the T-DNA.
            It indicates an insertion within the locus defined by BAC clone
            MRB17. Details on the protocols used for generation of the sequence
            are described in References 1-3. The sequences are generated at the
            MPI for Plant Breeding Research in the context of the GABI-Kat
            project. GABI-Kat is part of the German Plant Genomics program
            designated 'GABI'. Information on line availability can be found
            at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.

FEATURES             Location/Qualifiers
     1..165
     /organism="Arabidopsis thaliana"
     /mol_type="genomic DNA"
     /strain="Columbia 0"
     /db_xref="taxon:3702"

```

```

/clone="GK-523H09-020210"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Col-0"
/notes="PCR was performed on DNA from Arabidopsis thaliana
plants (T1) which were transformed with the T-DNA from
vector pAC161 (GenBank accession number: AJ537514). The
lines contain one or more T-DNA insertions. The DNA
fragment(s) resulting from the PCR were directly sequenced
to determine the genomic sequence flanking the insertion.
T-DNA derived sequences were removed."

ORIGIN
Query Match      100.0%; Score 11; DB 9; Length 165;
Best Local Similarity 100.0%; Pred. No. 8.2e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGCA 11
        |||||
Db      7 TAACTATATGCA 17

RESULT 75
AZ112443
LOCUS      AZ112443
DEFINITION RPCI-23-449D1.TJ RPCI-23 Mus musculus genomic clone RPCI-23-449D1,
            genomic survey sequence.
ACCESSION  AZ112443
VERSION     AZ112443.1 GI:7771343
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 166)
            Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S.,
            Akinret, B., Levine, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de
            Jong, P. and Fraser, C.M.
            Mouse BAC End Sequences from Library RPCI-23
            Unpublished (1999)
            Other GSSs: RPCI-23-449D1.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac/ends/mouse/bac_end_intro.html
            Plate: 449 row: D column: 1
            Seq primer: SP6
            Class: BAC ends.

FEATURES             Location/Qualifiers
     1..166
     /organism="Mus musculus"
     /mol_type="genomic DNA"
     /strain="C57BL/6J"
     /db_xref="taxon:10090"
     /clone="RPCI-23-449D1"
     /sex="Female"
     /lab_host="DH10B"
     /clone_lib="RPCI-23"
     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
            EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
            brain genomic DNA was isolated and partially digested
            with a combination of EcoRI and EcoRI Methyase. Size
            selected DNA was cloned into the pBACe3.6 vector at the
            EcoRI sites. The ligation products were transformed into
            DH10B electrocompetent cells (BRL Life Technologies)."

```

ORIGIN

Query Match 100.0%; Score 11; DB 8; Length 166;
 Best Local Similarity 100.0%; Pred. NO. 8.2e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 118 TAACTATATGA 128

Search completed: August 30, 2005, 12:50:57
 Job time : 45.5461 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 01:39:19 ; Search time 6.2234 Seconds
(without alignments)
10463.269 Million cell updates/sec

Title: US-09-983-000a-1_COPY_1262_1272

Perfect score: 11

Sequence: 1 taactatatga 11

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	100.0	25	9 ACI83570	AcI83570 Human mic
2	11	100.0	25	9 ACK12158	AcK12158 Human mic
3	11	100.0	145	2 AAK33881	AaK33881 Sequence
4	11	100.0	158	3 AAC11657	AaC11657 Human sec
5	11	100.0	183	12 ADL85349	AdL85349 DNA up-re
6	11	100.0	183	12 ADL85348	AdL85348 DNA up-re
7	11	100.0	217	8 ABX53848	AbX53848 Bovine ES
8	11	100.0	220	3 AAC28530	AaC28530 Human sec
9	11	100.0	228	3 AAC54673	AaC54673 Arabidops
10	11	100.0	235	13 ACN47121	AcN47121 Cotton pr
11	11	100.0	239	7 ADS68360	AdS68360 Corn seed
12	11	100.0	255	3 AAC08533	AaC08533 Human sec
13	11	100.0	268	5 AAC161531	AaC161531 Soybean 2
14	11	100.0	284	4 ABA07302	AbA07302 Human pan
15	11	100.0	284	4 ABA07305	AbA07305 Human pan
16	11	100.0	284	4 AAK90430	AaK90430 Human dig
17	11	100.0	284	4 AAK90433	AaK90433 Human dig
18	11	100.0	284	4 AAK74694	AaK74694 Human imm
19	11	100.0	298	2 AAT20828	AtA20828 Human gen
20	11	100.0	325	8 ABZ17767	AbZ17767 S2 subtra

C	21	11	100.0	329	4	AAS32888	Aa32888 Human gen
C	22	11	100.0	333	9	ADA30131	AdA30131 DNA encod
C	23	11	100.0	336	10	ADP02797	AdP02797 Bacterial
C	24	11	100.0	346	5	ABV60195	AbV60195 Human pro
C	25	11	100.0	347	4	AAS38386	Aa38386 Novel hum
C	26	11	100.0	348	6	ABL84127	AbL84127 Human ova
C	27	11	100.0	348	10	ADP00709	AdP00709 Bacterial
C	28	11	100.0	351	10	ADH83614	AdH83614 Enterococ
C	29	11	100.0	360	4	AAK61214	AaK61214 Human imm
C	30	11	100.0	370	5	ABV17250	AbV17250 Human pro
C	31	11	100.0	384	9	ACH30883	AcH30883 Human bon
C	32	11	100.0	392	6	ABQ57583	AbQ57583 Human col
C	33	11	100.0	395	4	AAH70968	AaH70968 Human cer
C	34	11	100.0	397	4	AAK89394	AaK89394 Human dig
C	35	11	100.0	397	4	AAK89949	AaK89949 Human dig
C	36	11	100.0	397	4	AAK88190	AaK88190 Human dig
C	37	11	100.0	397	5	AA331909	Aa331909 Human liv
C	38	11	100.0	397	6	ABN90264	AbN90264 Human liv
C	39	11	100.0	397	10	ABX86567	AbX86567 Corn ear
C	40	11	100.0	397	11	ADJ15177	AdJ15177 Human liv
C	41	11	100.0	399	13	ACN47251	AcN47251 Cotton pr
C	42	11	100.0	400	4	AAI92400	AaI92400 Human pol
C	43	11	100.0	403	4	AAK59040	AaK59040 Human imm
C	44	11	100.0	404	4	AAK74370	AaK74370 Human imm
C	45	11	100.0	415	4	AAK78936	AaK78936 Human imm
C	46	11	100.0	418	5	ABV49710	AbV49710 Human pro
C	47	11	100.0	423	8	ABX46997	AbX46997 Bovine ES
C	48	11	100.0	425	6	ABN73189	AbN73189 Bovine em
C	49	11	100.0	439	4	AAI25340	AaI25340 Human bre
C	50	11	100.0	439	5	ABV47045	AbV47045 Human pro
C	51	11	100.0	440	12	ADP95418	AdP95418 Cotton ex
C	52	11	100.0	449	6	ABN61851	AbN61851 Human can
C	53	11	100.0	460	3	AAO93005	AaO93005 Human sec
C	54	11	100.0	464	5	ABV52390	AbV52390 Human pro
C	55	11	100.0	465	4	AAI10825	AaI10825 Probe #75
C	56	11	100.0	465	4	ABA52475	AbA52475 Human foe
C	57	11	100.0	465	4	AAI32083	AaI32083 Probe #76
C	58	11	100.0	465	4	ABA42054	AbA42054 Human bre
C	59	11	100.0	465	4	ABA22267	AbA22267 Probe #73
C	60	11	100.0	465	4	AAK26194	AaK26194 Human bon
C	61	11	100.0	465	4	AAK00742	AaK00742 Human bra
C	62	11	100.0	465	4	ABS25784	AbS25784 Human liv
C	63	11	100.0	465	5	AAI00750	AaI00750 Probe #74
C	64	11	100.0	465	6	ABS00780	AbS00780 Human gen
C	65	11	100.0	470	9	ACH24813	AcH24813 Human adu
C	66	11	100.0	471	9	ACH26788	AcH26788 Human adu
C	67	11	100.0	476	4	AAI32986	AaI32986 Probe #16
C	68	11	100.0	476	4	ABA42958	AbA42958 Human bre
C	69	11	100.0	479	9	ACH23189	AcH23189 Human adu
C	70	11	100.0	479	10	ABX60813	AbX60813 Arabidops
C	71	11	100.0	482	5	ABV50770	AbV50770 Human pro
C	72	11	100.0	484	4	AAK89948	AaK89948 Human dig
C	73	11	100.0	484	4	AAK89391	AaK89391 Human dig
C	74	11	100.0	484	5	AA331906	Aa331906 Human liv
C	75	11	100.0	484	6	ABN90261	AbN90261 Human liv
C	76	11	100.0	484	11	ADJ15174	AdJ15174 Human liv
C	77	11	100.0	493	9	ACH32835	AcH32835 Human end
C	78	11	100.0	494	4	AAK87472	AaK87472 Human imm
C	79	11	100.0	498	5	ABV46111	AbV46111 Human pro
C	80	11	100.0	498	6	ABQ60129	AbQ60129 Human col
C	81	11	100.0	500	13	ADR61691	AdR61691 Cotton CD
C	82	11	100.0	501	13	ACN45788	AcN45788 Cotton pr
C	83	11	100.0	507	5	ABV55920	AbV55920 Human pro
C	84	11	100.0	511	4	AAI16497	AaI16497 Human bre
C	85	11	100.0	513	5	ABA12644	AbA12644 Human ner
C	86	11	100.0	521	12	ACH79220	AcH79220 Human gen
C	87	11	100.0	530	10	ADB50460	AdB50460 Primary r
C	88	11	100.0	532	5	ABV50729	AbV50729 Human pro
C	89	11	100.0	545	4	AAI11756	AaI11756 Human bre
C	90	11	100.0	545	4	AAK64404	AaK64404 Human imm
C	91	11	100.0	546	10	ADK52629	AdK52629 Plant DNA
C	92	11	100.0	548	8	ABX98484	AbX98484 Rice albu
C	93	11	100.0	550	3	AAC94956	AaC94956 Cat flea

c 94 11 100.0 552 5 ABA14741 Aba14741 Human ner
 c 95 11 100.0 555 4 ABA59777 ABA59777 Human foe
 c 96 11 100.0 555 4 AAI39648 AAI39648 Probe #83
 c 97 11 100.0 555 4 ABA28267 ABA28267 Probe #67
 c 98 11 100.0 555 4 AAK33922 AAK33922 Human bon
 c 99 11 100.0 555 4 AAK08048 AAK08048 Human bra
 c 100 11 100.0 555 4 ABS33728 ABS33728 Human liv

ALIGNMENTS

RESULT 1

ACI83570
 ID ACI83570 standard; DNA; 25 BP.

XX AC ACI83570;

XX 14-OCT-2003 (first entry)

XX Human microarray DNA oligonucleotide SEQ ID NO 83561.

XX EST; ss; probe; expressed sequence tag; microarray; gene expression;
 KW genetic variation; biallelic marker; polymorphism; human;
 KW cross-species comparison.

XX Homo sapiens.

XX US2003104410-A1.

XX 05-JUN-2003.

XX 15-MAR-2002; 2002US-00098263.

XX 16-MAR-2001; 2001US-0276759P.

XX (AFFY-) AFFYMETRIX INC.

XX Mittmann MP;

XX WPI; 2003-567953/53.

XX New array of nucleic acid probes, useful for in situ hybridization, in
 PT Southern, Northern or dot-blot hybridization to identify or detect the
 PT sequence or specific mutations of any gene.

XX Claim 1; SEQ ID NO 83561; 9pp; English.

XX The invention discloses a microarray comprising a plurality of nucleic
 CC acid probes including one of 2,018,500 fully defined sequences, or its
 CC perfect match, perfect mismatch, antisense match or antisense mismatch.
 CC Also disclosed is a method of gene expression analysis. The array is used
 CC in monitoring gene expression levels by hybridisation to a DNA library,
 CC in analysis of genetic variation or in hybridisation of tag-labelled
 CC compounds. The nucleic acid probes are specifically designed for analysis
 CC of at least one target sequence. The method of analysis comprises
 CC hybridising at least one or more nucleic acids to at least two or more
 CC nucleic acid probes and detecting the hybridisation. The nucleic acid
 CC probes are attached to a solid support. The analysis comprises monitoring
 CC gene expression levels, identifying biallelic markers or polymorphisms,
 CC or family members of a gene and a cross-species comparison. Each of the
 CC nucleic acids further comprises a tag sequence. The array of nucleic acid
 CC probes is useful in situ hybridisation, in Southern, Northern or dot-
 CC blot hybridisation to identify or detect the sequence or specific
 CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
 CC primer extensions or in screening cDNA or genomic libraries or subclones
 CC for additional subclones containing segments of DNA that have been
 CC isolated and previously sequenced. The sequence presented is one of the
 CC nucleic acid probes incorporated in the microarray. Note: The sequence
 CC data for this patent can also be obtained in electronic format directly
 CC from USPTO at seqdata.uspto.gov/sequence.html

XX Sequence 25 BP; 12 A; 4 C; 4 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 11; DB 9; Length 25;
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11

DB 7 TAACTATATGA 17

RESULT 2

ACK12158

ID ACK12158 standard; DNA; 25 BP.

XX AC ACK12158;

XX 14-OCT-2003 (first entry)

XX Human microarray DNA oligonucleotide SEQ ID NO 112139.

XX EST; ss; probe; expressed sequence tag; microarray; gene expression;
 KW genetic variation; biallelic marker; polymorphism; human;
 KW cross-species comparison.

XX Homo sapiens.

XX US2003104410-A1.

XX 05-JUN-2003.

XX 15-MAR-2002; 2002US-00098263.

XX 16-MAR-2001; 2001US-0276759P.

XX (AFFY-) AFFYMETRIX INC.

XX Mittmann MP;

XX WPI; 2003-567953/53.

XX New array of nucleic acid probes, useful for in situ hybridization, in
 PT Southern, Northern or dot-blot hybridization to identify or detect the
 PT sequence or specific mutations of any gene.

XX Claim 1; SEQ ID NO 112139; 9pp; English.

XX The invention discloses a microarray comprising a plurality of nucleic
 CC acid probes including one of 2,018,500 fully defined sequences, or its
 CC perfect match, perfect mismatch, antisense match or antisense mismatch.
 CC Also disclosed is a method of gene expression analysis. The array is used
 CC in monitoring gene expression levels by hybridisation to a DNA library,
 CC in analysis of genetic variation or in hybridisation of tag-labelled
 CC compounds. The nucleic acid probes are specifically designed for analysis
 CC of at least one target sequence. The method of analysis comprises
 CC hybridising at least one or more nucleic acids to at least two or more
 CC nucleic acid probes and detecting the hybridisation. The nucleic acid
 CC probes are attached to a solid support. The analysis comprises monitoring
 CC gene expression levels, identifying biallelic markers or polymorphisms,
 CC or family members of a gene and a cross-species comparison. Each of the
 CC nucleic acids further comprises a tag sequence. The array of nucleic acid
 CC probes is useful in situ hybridisation, in Southern, Northern or dot-
 CC blot hybridisation to identify or detect the sequence or specific
 CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
 CC primer extensions or in screening cDNA or genomic libraries or subclones
 CC for additional subclones containing segments of DNA that have been
 CC isolated and previously sequenced. The sequence presented is one of the
 CC nucleic acid probes incorporated in the microarray. Note: The sequence
 CC data for this patent can also be obtained in electronic format directly
 CC from USPTO at seqdata.uspto.gov/sequence.html

XX Sequence 25 BP; 12 A; 4 C; 4 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 11; DB 9; Length 25;

ID	AAC11657 standard; cDNA; 158 BP.
XX	
AC	AAC11657;
AC	
XX	
DT	06-OCT-2000 (first entry)
XX	
DE	Human secreted protein 5' EST, SEQ ID NO: 15732.
XX	
KW	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW	gene therapy; chromosome mapping; ss.
XX	
OS	Homo sapiens.
XX	
PN	EP1033401-A2.
XX	
PD	06-SEP-2000.
XX	
PF	21-FEB-2000; 2000EP-00200610.
XX	
PR	26-FEB-1999; 99US-0122487P.
XX	
PA	(GEST) GENSET.
XX	
PI	Dumas Milne Edwards J, Duclert A, Giordano J;
XX	
DR	WPI; 2000-500381/45.
XX	
PT	New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT	obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT	diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX	
PS	Claim 1; SEQ ID NO 15732; 71pp + Sequence Listing; English.
XX	
CC	The present sequence is one of a large number of 5' ESTs derived from
CC	mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC	identified within the present sequence. The 5' ESTs were prepared from
CC	total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC	sequences usually correspond mainly to the 3' untranslated region (UTR)
CC	of the mRNA because they are often obtained from oligo-dT primed cDNA
CC	libraries. Such ESTs are not well suited for isolating cDNA sequences
CC	derived from the 5' ends of mRNAs and even in those cases where longer
CC	cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC	ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC	to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC	diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC	They are used to obtain upstream regulatory sequences and to design
CC	expression and secretion vectors
XX	
SQ	Sequence 158 BP; 61 A; 34 C; 21 G; 41 T; 0 U; 1 Other;
	Query Match 100.0%; Score 11; DB 3; Length 158;
	Best Local Similarity 100.0%; Pred. No. 1.8e+03;
	Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps
Qy	1 TAACTATATGA 11
Db	27 TAACTATATGA 37
RESULT 5	
ADL85349/c	
ID	ADL85349 standard; DNA; 183 BP.
XX	
AC	ADL85349;
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	DNA up-regulated in murine common lymphoid myeloid cells SeqID 1742.
XX	
KW	gene potential; multi-lineage; cell commitment; haematopoietic stem cell
KW	HSC; multipotent progenitor; MPP; common lymphoid progenitor; CLP;
KW	common myeloid progenitor; CMP; bone marrow stem cell; mouse; murine; ds
XX	

```

OS Mus sp.
XX WO2003093445-A2.
XX
XX 13-NOV-2003,
XX
XX 05-MAY-2003) 2003WO-US014114.
XX
XX 03-MAY-2002) 2002US-0377383P.
XX
XX (STOW-) STOWERS INST MEDICAL RES.
XX
XX Li L;
XX WPI; 2004-042656/02.
XX
XX Classifying an unknown multi-lineage affiliated gene comprises isolating
XX expressed nucleic acid sequences from the discrete cell sub-populations.
XX
XX Claim 9; SEQ ID NO 1742; 123pp; English.
XX
XX This invention relates to a novel method for predicting gene potential by
XX associating nucleic acid sequences of unknown function with particular
XX sub-population profiles. Specifically, it refers to classifying an
XX unknown multi-lineage affiliated gene by collecting hybridisation data to
XX develop a gene expression map, in order to determine the discrete sub-
XX population where it is expressed. The present invention describes methods
XX for predicting the lineage commitment of genes associated with the self-
XX renewing haematopoietic (blood) stem cells (HSCs), as well as the non-
XX self renewing multipotent progenitors (MPPs), common lymphoid progenitors
XX (CLPs) and common myeloid progenitors (CMPs), which are collectively
XX referred to as bone marrow stem cells populations. As such, these methods
XX can be used to identify associated multi-lineage affiliated genes and
XX hence the underlying molecular mechanisms in physiological haematopoietic
XX development. This polynucleotide sequence is DNA associated with a murine
XX CMP sub population of cells of the invention.
XX
XX Sequence 183 BP; 60 A; 33 C; 29 G; 61 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 11; DB 12; Length 183;
XX Best Local Similarity 100.0%; Pred. No. 1.8e+03;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 TAACTATATGCA 11
XX |||||
XX 154 TAACTATATGCA 144
XX
XX RESULT 6
XX ADL85348/c
XX ID ADL85348 standard; DNA; 183 BP.
XX
XX AC ADL85348;
XX
XX DT 20-MAY-2004 (first entry)
XX
XX DE DNA up-regulated in murine common lymphoid myeloid cells SeqID 1741.
XX
XX gene potential; multi-lineage; cell commitment; haematopoietic stem cell;
XX HSC; multipotent progenitor; MPP; common lymphoid progenitor; CLP;
XX common myeloid progenitor; CMP; bone marrow stem cell; mouse; murine; ds.
XX
XX Mus sp.
XX WO2003093445-A2.
XX
XX 13-NOV-2003.
XX
XX 05-MAY-2003; 2003WO-US014114.
XX
XX 03-MAY-2002; 2002US-0377383P.
XX
XX (STOW-) STOWERS INST MEDICAL RES.
XX
XX Purified nucleic acid molecules, useful for genome mapping, gene
XX identification and analysis, cattle breeding or preparation of constructs

```

```

XX Li L;
XX WPI; 2004-022656/02.
XX
XX Classifying an unknown multi-lineage affiliated gene comprises isolating
XX expressed nucleic acid sequences from the discrete cell sub-populations.
XX
XX Claim 9; SEQ ID NO 1741; 123pp; English.
XX
XX This invention relates to a novel method for predicting gene potential by
XX associating nucleic acid sequences of unknown function with particular
XX sub-population profiles. Specifically, it refers to classifying an
XX unknown multi-lineage affiliated gene by collecting hybridisation data to
XX develop a gene expression map, in order to determine the discrete sub-
XX population where it is expressed. The present invention describes methods
XX for predicting the lineage commitment of genes associated with the self-
XX renewing haematopoietic (blood) stem cells (HSCs), as well as the non-
XX self renewing multipotent progenitors (MPPs), common lymphoid progenitors
XX (CLPs) and common myeloid progenitors (CMPs), which are collectively
XX referred to as bone marrow stem cells populations. As such, these methods
XX can be used to identify associated multi-lineage affiliated genes and
XX hence the underlying molecular mechanisms in physiological haematopoietic
XX development. This polynucleotide sequence is DNA associated with a murine
XX CMP sub population of cells of the invention.
XX
XX Sequence 183 BP; 60 A; 33 C; 29 G; 61 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 11; DB 12; Length 183;
XX Best Local Similarity 100.0%; Pred. No. 1.8e+03;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 TAACTATATGCA 11
XX |||||
XX 154 TAACTATATGCA 144
XX
XX RESULT 7
XX ABX53848/c
XX ID ABX53848 standard; cDNA; 217 BP.
XX
XX AC ABX53848;
XX
XX DT 25-FEB-2003 (first entry)
XX
XX DE Bovine EST associated with lactation/muscle/fat deposition #3777.
XX
XX Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
XX muscle deposition; fat deposition; genome mapping; gene identification;
XX gene analysis; cattle breeding.
XX
XX Bos Taurus.
XX
XX US2002137160-A1.
XX
XX 26-SEP-2002.
XX
XX 26-OCT-2001; 2001US-00983965.
XX
XX 17-DEC-1998; 98US-0113678P.
XX
XX 15-DEC-1999; 99US-00465231.
XX
XX (BYAT/) BYATT J C.
XX (MATH/) MATHIALAGAN N.
XX (TAON/) TAO N.
XX (WARR/) WARREN W C.
XX
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-102386/09.
XX
XX Purified nucleic acid molecules, useful for genome mapping, gene
XX identification and analysis, cattle breeding or preparation of constructs

```


PT for cattle gene expression and genetically improved cattle.

XX Claim 2; SEQ ID NO 3777; 38pp; English.

XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
CC nucleic acid molecule comprising any of 5912 nucleotide sequences,
CC appearing as ABX50072-ABX55983, or complements of them. Also included are
CC : (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non-translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 5912 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the 5912
CC bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The present
CC sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?docID=20020137160
XX

SQ Sequence 217 BP; 53 A; 56 C; 42 G; 66 T; 0 U; 0 Other;

Query Match 100.0%; Score 11; DB 8; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
|||||
DB 97 TAACTATATGCA 87

RESULT 8

AAC28530
ID AAC28530 standard; cDNA; 220 BP.

AC AAC28530;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 32605.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.

XX (GEST) GENBET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for

PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 1; SEQ ID NO 32605; 71pp + Sequence Listing; English.

XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX

SQ Sequence 220 BP; 74 A; 28 C; 46 G; 69 T; 0 U; 3 Other;

Query Match 100.0%; Score 11; DB 3; Length 220;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
|||||
DB 201 TAACTATATGCA 211

RESULT 9

AAC54673
ID AAC54673 standard; DNA; 228 BP.

AC AAC54673;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 78678.

XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0121180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130445P.

XX 23-APR-1999; 99US-0130510P.

XX 23-APR-1999; 99US-0130891P.

XX 28-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 30-APR-1999; 99US-0132407P.

XX 04-MAY-1999; 99US-0132484P.

XX 06-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135114P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136332P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137282P.
PR 03-JUN-1999; 99US-0137520P.
PR 04-JUN-1999; 99US-0137724P.
PR 07-JUN-1999; 99US-0138094P.
PR 08-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154799P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.

PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161932P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 11; DB 3; Length 228;
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
 Db 36 TAACTATATGCA 46

RESULT 10
 ACN47121
 ID ACN47121 standard; cDNA; 235 BP.
 XX AC
 AC ACN47121;
 XX AC
 DT 02-DEC-2004. (first entry)
 XX DT
 DE Cotton primed seed EST Clone ID: L1B3825-007-Q1-N6-C8, SEQ:1902.
 XX DE
 XX Cotton; plant; EST: expressed sequence tag; transgenic plant; seed;
 KW variety DP508; library L1B3825; molecular tag; molecular marker;
 KW genetic mapping; molecular mapping; seed germination; plant growth;
 KW plant quality; plant yield; plant breeding; tissue printing; ss.
 XX KW
 OS Gossypium hirsutum.
 XX OS
 XX US2004123340-A1.
 XX PN
 PD 24-JUN-2004.
 XX PD
 XX 12-DEC-2001; 2001US-00021323.
 XX PF
 XX 14-DEC-2000; 2000US-0255619P.
 XX PR
 PA (DEIK/) DEIKMAN J.
 PA (FENG/) FENG P C C.
 PA (FINC/) FINCHER K L.
 PA (ZIEG/) ZIEGLER T B.
 XX PA
 XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;
 XX PI
 XX WPI; 2004-479808/45.
 XX DR
 XX New isolated nucleic acid molecule that encodes a plant protein or its
 PT fragment, useful for isolating a variety of agronomically significant
 PT genes associated with plant growth, quality or yield, and as molecular
 PT tags to map genes.
 XX PT
 XX Claim 1; SEQ ID NO 1902; 34pp; English.
 XX PS
 XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
 CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
 CC from primed or non-primed seeds from variety DP508, mature seeds from
 CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
 CC tissue, developing fibres, carpel walls and septa from variety
 CC Nucleon33B. The invention also relates to substantially purified
 CC proteins or their fragments encoded by nucleic acid molecules of the
 CC invention, and to transformed plants having a nucleic acid construct
 CC comprising a nucleic acid of the invention. The cotton ESTs are useful as

molecular tags to isolate genetic regions, to isolate genes, to map
 genes, to determine gene function and to determining whether genes are
 members of a particular gene family. The nucleic acid molecules may be
 used for isolating a variety of agronomically significant genes
 associated with plant growth, quality, yield, and could also serve as
 links in metabolic and catabolic pathways. The nucleic acid molecules are
 also useful for identifying genes important in initiating and maintaining
 seed germination or that may be used to mitigate stresses encountered
 during seed germination. The ESTs additionally enable the acquisition of
 promoters and cis-regulatory elements which will be useful to express
 agronomically significant genes in these tissues and/or other tissues,
 and also permits the acquisition of molecular markers useful in breeding
 schemes, genetic and molecular mapping, and in cloning of agronomically
 significant genes. The nucleic acid molecules are further useful for
 detecting the expression level or pattern of a protein by tissue printing. The
 present sequence represents a specifically claimed EST isolated from a
 cotton variety DP508 primed seed cDNA library (L1B3825). The sequence
 data for this patent did not form part of the printed specification, but
 was obtained in electronic format directly from the US patent office at
 seqdata.uspto.gov/sequence.html?DocID=US20040123340

Query Match 100.0%; Score 11; DB 13; Length 235;
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
 Db 95 TAACTATATGCA 105

RESULT 11
 ADS68360
 ID ADS68360 standard; cDNA; 239 BP.
 XX AC
 AC ADS68360;
 XX AC
 DT 18-NOV-2004 (first entry)
 XX DT
 DE Corn seedling-derived polynucleotide (cpds), SEQ ID 3376.
 XX DE
 XX Corn; seedling-derived polynucleotide; cdp; ss; plant; seed growth;
 KW seed development; disease resistance; insect infestation; fungal disease;
 KW bacterial infection; Goss' Bacterial Wilt; blight;
 KW Stewart's bacterial wilt; Holcusa spot; bacterial leaf blight; leaf spot;
 KW bacterial stripe; maize dwarf mosaic virus infection;
 KW environmental stress; water stress; pH stress; temperature stress;
 KW pollution; injury; pesticide.
 XX KW
 OS Zea mays.
 XX OS
 XX US2003237110-A9.
 XX PN
 XX 25-DEC-2003.
 XX PD
 XX 06-AUG-2001; 2001US-00923876.
 XX PF
 XX 12-MAY-1998; 98US-0085331P.
 XX PR
 XX 21-APR-1999; 99US-00298329.
 XX PR
 XX (INCY-) INCYTE PHARM INC.
 XX PA
 XX Lalgudi RV, Ito LY, Sherman BK;
 XX PI
 XX WPI; 2002-195165/25.
 XX DR
 XX New corn seedling-derived polynucleotides and polypeptides, useful in
 PT identifying and altering desired characteristics associated with growth
 PT and development, disease resistance, environmental adaptability, quality
 PT and yield.
 XX PT

PS Claim 1; SEQ ID NO 3376; 33pp; English.

XX

CC The invention relates to a corn seedling-derived polynucleotide (cdp) selected from ADS64985-ADS71316, or their complements and fragments. Also included are a composition for the detection of altered expression of a cdp (comprising a polynucleotide selected from ADS64985-ADS71316), a method of detecting a polynucleotide in a biological sample using a cdp, a method for using oligomers (and amplification) to recover a regulatory element from a DNA library using oligomers designed against a cdp, a seedling specific regulatory element that regulates the expression of a cdp, an expression vector containing a cdp or regulatory element, a plant transformed with the vector, a host cell containing the vector (and expressing a corn seedling derived protein, CDP), an anti-CDP antibody, identifying a compound which binds a CDP and screening a plurality of compounds for binding to cdp polynucleotide. The cdp polynucleotides, proteins, vectors, cells and antibodies are useful for the identification, evaluation and alteration of seed growth and development, disease resistance (e.g. to insect infestation, fungal disease, bacterial infection, Goss' Bacterial Wilt, blight, Stewart's bacterial wilt, Holcous spot, bacterial leaf blight, leaf spot, bacterial stripe and maize dwarf mosaic virus infection) and resistance to environmental stress (e.g. water stress, pH stress, temperature stress, pollution, injury or pesticides). The present sequence is cdp cDNA sequence.

XX

SQ Sequence 239 BP; 52 A; 46 C; 51 G; 90 T; 0 U; 0 Other;

Query Match 100.0%; Score 11; DB 7; Length 239;
Best Local Similarity 100.0%; Pred. NO. 1.8e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TAACTATATGA 11
|||||
Db 29 TAACTATATGA 39

RESULT 12

AAC08533

ID AAC08533 standard; cDNA; 255 BP.

XX AAC08533;

XX

XX 06-OCT-2000 (first entry)

XX

DE Human secreted protein 5' EST, SEQ ID NO: 12608.

XX

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping; ss.

XX

OS Homo sapiens.

XX

XX EP1033401-A1.

XX

PD 06-SEP-2000.

XX

XX 21-FEB-2000; 2000EP-00200610.

XX

XX 26-FEB-1999; 99US-0122487P.

XX

XX (GEST) GENSET.

XX

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX

XX WPI; 2000-500381/45.

XX

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX

PS Claim 1; SEQ ID NO 12608; 71pp + Sequence Listing; English.

XX

CC The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. No ORF has yet been conclusively identified within the present sequence. The 5' ESTs were prepared from

CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors

XX

SQ Sequence 255 BP; 76 A; 53 C; 21 G; 105 T; 0 U; 0 Other;

Query Match 100.0%; Score 11; DB 3; Length 255;
Best Local Similarity 100.0%; Pred. NO. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TAACTATATGA 11
|||||
Db 196 TAACTATATGA 206

RESULT 13

AAI61531

ID AAI61531 standard; DNA; 268 BP.

XX AAI61531;

XX

XX 16-OCT-2001 (first entry)

XX

DE Soybean 240017 region G3 DNA, SEQ ID NO: 162.

XX

KW Soybean; antihelminthic; gene therapy; soybean cyst nematode; SCN;

KW SCN resistance; rhg1; Rhg4; SCN resistant allele; plant breeding;

KW 240017 region G3; 318013 region A3; 515002 region G2; ds.

XX

OS Glycine max.

XX

XX WO200151627-A2.

XX

XX 19-JUL-2001.

XX

XX 05-JAN-2001; 2001WO-US000552.

XX

XX 07-JAN-2000; 2000US-0174880P.

XX

XX (MONS) MONSANTO CO.

XX

XX Hauge BM, Wang ML, Parsons JD, Parnell LD;

XX

XX WPI; 2001-425872/45.

XX

XX New purified nucleic acid for producing a soybean plant having soybean cyst nematode resistance and for use in plant breeding programs.

XX

XX Claim 25; Page 1066; 1353pp; English.

XX

CC The invention relates to nucleic acid molecules from regions of the soybean genome which are associated with soybean cyst nematode (SCN) resistance. The nucleic acids are used to transform plants, and can produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele. The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes of soybean plants and for introgressing SCN resistance or partial SCN resistance into soybean plants. They can also be used in plant breeding programmes. The invention also relates to proteins encoded by such nucleic acid molecules, as well as antibodies capable of recognising these proteins. The present sequence is a nucleic acid molecule provided in the specification

XX

SQ Sequence 268 BP; 123 A; 41 C; 25 G; 79 T; 0 U; 0 Other;

Query Match 100.0%; Score 11; DB 5; Length 268;

Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TAACTATATGA 11
Db 40 TAACTATATGA 50

RESULT 14

ABA07302/C
ID ABA07302 standard; DNA; 284 BP.

AC ABA07302;

XX
DT 14-JAN-2002 (first entry)

XX Human pancreatic cancer related genomic DNA, SEQ ID NO: 621.

XX Human; cytostatic; antidiabetic; antiinflammatory; gastric; osteopathic;
KW antihormone; antiulcer; thyroid-active; gene therapy; antisense therapy;
KW pancreatic cancer antigen inhibitor; pancreatic cancer; pancreatitis;
KW diabetes; endocrine disorder; acromegaly; hyperthyroidism;
KW gastrointestinal disorder; Crohn's disease; duodenal ulcer; ds.

OS Homo sapiens.

XX WO200155206-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001353.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217486P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234597P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-457717/49.
XX
XX Isolated pancreatic cancer polypeptide for treating, preventing and/ or
XX prognosing disorders related to the pancreas including pancreatic cancers
XX and also for testing and detection e.g. diagnosis.
XX
XX Disclosure; SEQ ID NO 621; 537pp; English.
XX
XX The invention relates to an isolated polypeptide comprising an amino acid
XX sequence at least 90% identical to 188 amino acid sequences fully defined
XX in the specification and encoded by 188 cDNA clones fully defined in the
XX specification. The invention also relates to a fragment having biological
XX activity, a domain, an epitope, full length protein, variant, allelic
XX variant or a species homologue of the fully defined sequence. The
XX polynucleotide and polypeptide are useful for treating, preventing and/or
XX prognosing disorders related to the pancreas including pancreatic cancer,
XX pancreatitis, diabetes, endocrine disorders such as acromegaly or
XX hyperthyroidism, and gastrointestinal disorders such as Crohn's disease
XX and duodenal ulcers. The present sequence encodes a pancreatic cancer-
XX related polypeptide of the invention
XX
XX Sequence 284 BP; 113 A; 43 C; 44 G; 84 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 11; DB 4; Length 284;
XX Best Local Similarity 100.0%; Pred. No. 1.7e+03;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 TAACTATATCA 11
XX |||||
XX 163 TAACTATATCA 153
XX
XX
XX RESULT 15
XX ABA07305/c
XX ID ABA07305 standard; DNA; 284 BP.
XX AC ABA07305;
XX
XX 14-JAN-2002 (first entry)
XX
XX Human pancreatic cancer related genomic DNA, SEQ ID NO: 624.
XX
XX Human; cytostatic; antidiabetic; antiinflammatory; gastric; osteopathic;
XX antihormone; antiulcer; thyroid-active; gene therapy; antisense therapy;
XX pancreatic cancer antigen inhibitor; pancreatic cancer; pancreatitis;
XX diabetes; endocrine disorder; acromegaly; hyperthyroidism;

KW gastrointestinal disorder; Crohn's disease; duodenal ulcer; ds.
XX Homo sapiens.
OS WO200155206-A1.
PN
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001353.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 12-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0232403P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 21-SEP-2000; 2000US-0233065P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.

```
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249237P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0254937P.
PR 05-JAN-2001; 2001US-0255967P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-457717/49.
XX Isolated pancreatic cancer polypeptide for treating, preventing and/or
XX prognosing disorders related to the pancreas including pancreatic cancers
XX and also for testing and detection e.g. diagnosis.
XX Disclosure; SEQ ID NO 624; 537pp; English.
XX The invention relates to an isolated polypeptide comprising an amino acid
XX sequence at least 90% identical to 188 amino acid sequences fully defined
XX in the specification and encoded by 188 cDNA clones fully defined in the
XX specification. The invention also relates to a fragment having biological
XX activity, a domain, an epitope, full length protein, variant, allelic
XX variant or a species homologue of the fully defined sequence. The
XX polynucleotide and polypeptide are useful for treating, preventing and/or
XX prognosing disorders related to the pancreas including pancreatic cancer,
XX pancreatitis, diabetes, endocrine disorders such as acromegaly or
XX hyperthyroidism, and gastrointestinal disorders such as Crohn's disease
XX and duodenal ulcers. The present sequence encodes a pancreatic cancer-
XX related polypeptide of the invention
XX Sequence 284 BP; 113 A; 43 C; 44 G; 84 T; 0 U; 0 Other;
Query Match 100.0%; Score 11; DB 4; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TRACTATATGA 11
DB 163 TRACTATATGA 153
RESULT 16
AAK90430/C
ID AAK90430 standard; DNA; 284 BP.
XX
XX AAK90430;
XX
XX 05-NOV-2001 (first entry)
XX Human digestive system antigen genomic sequence SEQ ID NO: 4006.
XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
XX ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
XX digestive system disorder; Meckel's diverticulum; ds.
XX Homo sapiens.
XX
XX WO200155314-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001324.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
```

PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218230P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-502630/55.
XX
XX PT Polynucleotides encoding digestive system antigens, useful for
XX PT diagnosing, treating, preventing and/or prognosing disorders of the
XX PT digestive system, particularly cancer and cancer metastases.
XX
XX PS Disclosure; SEQ ID NO 4006; 986pp; English.
XX
XX CC The present invention provides the protein and coding sequences of a
XX CC number of human digestive system antigens. These can be used in the
XX CC diagnosis, treatment and prevention of digestive system disorders,
XX CC including cancer, Meckel's diverticulum, bacterial or parasitic
XX CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
XX CC ulcerative colitis. The present sequence is a genomic DNA fragment
XX CC encoding a digestive system antigen of the invention
SQ Sequence 284 BP; 113 A; 43 C; 44 G; 84 T; 0 U; 0 Other;
Query Match 100.0%; Score 11; DB 4; Length 284;


```
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TAACTATATGA 11
Db      163 TAACTATATGA 153

RESULT 17
ID AAK90433/c
ID AAK90433 standard; DNA; 284 BP.
XX
AC AAK90433;
XX
DT DT
XX
DE DE
XX
KW Human digestive system antigen genomic sequence SEQ ID NO: 4009.
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ds.
XX
OS Homo sapiens.
XX
FN WO200155314-A2.
XX
PD PD
XX
PF 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001324.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-01981123P.
PR 07-JUN-2000; 2000US-0205515P.
PR 19-MAY-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
```

```

PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-502630/55.
XX
XX Polynucleotides encoding digestive system antigens, useful for
XX diagnosing, treating, preventing and/or prognosing disorders of the
XX digestive system, particularly cancer and cancer metastases.
XX
XX Disclosure; SEQ ID NO 4009; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human digestive system antigens. These can be used in the
XX diagnosis, treatment and prevention of digestive system disorders,
XX including cancer, Meckel's diverticulum, bacterial or parasitic
XX infections, appendicitis, Hirschsprung's disease, chronic colitis or
XX ulcerative colitis. The present sequence is a genomic DNA fragment
XX encoding a digestive system antigen of the invention
XX
XX Sequence 284 BP; 113 A; 43 C; 44 G; 84 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 11; DB 4; Length 284;
XX Best Local Similarity 100.0%; Pred. No. 1.7e+03;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 TAACTATATGCA 11
XX |||||
XX Db 163 TAACTATATGCA 153
XX
XX RESULT 18
XX AAK74694
XX ID AAK74694 standard; DNA; 284 BP.
XX
XX AC AAK74694;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29506.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX

```

```

PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.

```

02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239935P.
13-OCT-2000; 2000US-0239937P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241221P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241786P.
20-OCT-2000; 2000US-0241787P.
20-OCT-2000; 2000US-0241808P.
20-OCT-2000; 2000US-0241809P.
20-OCT-2000; 2000US-0241826P.
01-NOV-2000; 2000US-0244617P.
08-NOV-2000; 2000US-0246474P.
08-NOV-2000; 2000US-0246475P.
08-NOV-2000; 2000US-0246476P.
08-NOV-2000; 2000US-0246477P.
08-NOV-2000; 2000US-0246478P.
08-NOV-2000; 2000US-0246523P.
08-NOV-2000; 2000US-0246524P.
08-NOV-2000; 2000US-0246525P.
08-NOV-2000; 2000US-0246526P.
08-NOV-2000; 2000US-0246527P.
08-NOV-2000; 2000US-0246528P.
08-NOV-2000; 2000US-0246532P.
08-NOV-2000; 2000US-0246609P.
08-NOV-2000; 2000US-0246610P.
08-NOV-2000; 2000US-0246611P.
08-NOV-2000; 2000US-0246613P.
17-NOV-2000; 2000US-0249207P.
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249244P.
17-NOV-2000; 2000US-0249245P.
17-NOV-2000; 2000US-0249264P.
17-NOV-2000; 2000US-0249265P.
17-NOV-2000; 2000US-0249293P.
17-NOV-2000; 2000US-0249299P.
17-NOV-2000; 2000US-0249300P.
01-DEC-2000; 2000US-0250391P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
06-DEC-2000; 2000US-0256719P.
06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254097P.
05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-483426/52.
Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and metastasis.

PS Disclosure; SEQ ID NO 29506; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 284 BP; 124 A; 30 C; 52 G; 78 T; 0 U; 0 Other;
Query Match 100.0%; Score 11; DB 4; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAACTATATGA 11
Db 83 TAACTATATGA 93
|||||
RESULT 19
AAT20828
ID AAT20828 standard; cDNA to mRNA; 298 BP.
XX AAT20828;
XX 10-JUL-1996 (first entry)
XX Human gene signature HUMGS02077.
XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX Homo sapiens.
XX WO9514772-A1.
PN 01-JUN-1995.
XX 11-NOV-1994; 94WO-JP001916.
XX 12-NOV-1993; 93JP-00355504.
XX (MATS/) MATSUBARA K.
XX (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
XX WPI; 1995-206931/27.
PT Single-stranded DNA for identifying gene signatures - isolated from 3'-
PT directed human cDNA library that reflects relative abundance of corresp.
PT mRNA in specific human tissues.
XX Claim 1; Page 735; 2245pp; Japanese.
XX A single-stranded DNA (or its complementary strand or the corresp. double
CC stranded DNA) which comprises one of the 7837 "GS" sequences given in
CC AAT19001-T26837 and which is able to hybridise to part of human genomic
CC DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences were
CC obtained from 3'-directed cDNA libraries prepared from various human

CC tissues; synthesis of cDNA was initiated from the 3'-end of mRNA by using
 CC poly(7) as the sole primer. Since the 3'- untranslated sequence is unique
 CC to a particular mRNA species, almost all the 3'-oriented cDNAs hybridise
 CC with specific mRNAs. Each library is constructed so as to reflect
 CC accurately the relative abundance of different mRNAs in the particular
 CC tissue from which it was derived. The appearance frequency of a given GS
 CC in a cDNA library can be determined (esp. using primers and probes
 CC derived from the GS sequences) as a means of diagnosing abnormal cell
 CC function or for recognising different cell types
 XX

SQ Sequence 298 BP; 98 A; 47 C; 51 G; 101 T; 0 U; 1 Other;

Query Match 100.0%; Score 11; DB 2; Length 298;
 Best Local Similarity 100.0%; Pred. NO. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
 |||||
 DB 124 TAACTATATGA 134

RESULT 20
 ABZ17767/c
 ID ABZ17767 standard; cDNA; 325 BP.

XX AC ABZ17767;

XX DT 23-JAN-2003 (first entry)

XX DE S2 subtraction library cancer related clone SEQ ID NO:193.

XX KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
 KW immune response; virology; immunology; microbiology; molecular biology;
 KW recombinant DNA technology; Gene; ss.

XX OS Homo sapiens.

XX PN W0200278516-A2.

XX PD 10-OCT-2002;

XX PF 28-MAR-2002; 2002WO-US010421.

XX PR 30-MAR-2001; 2001US-0280255P.

XX PR 28-AUG-2001; 2001US-0315563P.

XX PR 09-JAN-2002; 2002US-0347313P.

XX PA (CORI-) CORIXA CORP.

XX PI Wang T, Wang S, Bangur CS, Gaiger A;

XX DR WPI; 2003-098387/05.

XX PT New immunogenic polynucleotides or polypeptides useful for diagnosing,
 PT preventing and treating cancer expressing CT or CP mRNA antigens, and in
 PT virology, immunology, microbiology, molecular biology and recombinant DNA
 PT techniques.

XX PS Claim 1; SEQ ID NO 193; 207pp; English.

XX CC ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
 CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
 CC invention. (I) and (II) have cytostatic activity and can be used in gene
 CC therapy and vaccines. (I), (II), antibodies and compositions from the
 CC present invention are useful for diagnosing, preventing and treating
 CC cancer, which expresses CT or CP mRNA antigens. They are useful for
 CC stimulating immune response. They can also be useful in virology,
 CC immunology, microbiology, molecular biology and recombinant DNA
 CC techniques. N.B. The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 325 BP; 82 A; 56 C; 60 G; 126 T; 0 U; 1 Other;

Query Match 100.0%; Score 11; DB 8; Length 325;
 Best Local Similarity 100.0%; Pred. NO. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
 |||||
 DB 34 TAACTATATGA 24

RESULT 21

AAS32888/c

ID AAS32888 standard; DNA; 329 BP.

XX AC AAS32888;

XX DT 17-DEC-2001 (first entry)

XX DE Human genomic DNA for novel endocrine antigen, SEQ ID No 842.

XX KW Human; endocrine antigen; ds; cytostatic; antiinfertility; antidiabetic;
 KW thyroid-active; adrenal-active; androgenic; gastric; gene therapy;
 KW antisense-therapy; antibody; endocrine disorder; hormone imbalance;
 KW reproductive disorder; endocrine cancer; pancreatic disorder;
 KW diabetes mellitus; adrenal gland disorder; hirsutism; thyroid disorder;
 KW hyperthyroidism; hypothalamic disorder; vanishing testes syndrome.

XX OS Homo sapiens.

XX PN W0200155319-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001335.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 11-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225266P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225758P.

XX PR 14-AUG-2000; 2000US-0225759P.

XX PR 18-AUG-2000; 2000US-0226279P.

XX PR 22-AUG-2000; 2000US-0226681P.

XX PR 22-AUG-2000; 2000US-0226682P.

XX PR 22-AUG-2000; 2000US-0227182P.

XX PR 23-AUG-2000; 2000US-0227009P.

XX PR 30-AUG-2000; 2000US-0228924P.

XX PR 01-SEP-2000; 2000US-0229287P.

XX PR 01-SEP-2000; 2000US-0229343P.

PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0229437P.
PR	06-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231142P.
PR	08-SEP-2000;	2000US-0231143P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232080P.
PR	08-SEP-2000;	2000US-0232081P.
PR	12-SEP-2000;	2000US-02321968P.
PR	14-SEP-2000;	2000US-0232397P.
PR	14-SEP-2000;	2000US-0232398P.
PR	14-SEP-2000;	2000US-0232399P.
PR	14-SEP-2000;	2000US-0232400P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-02323063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233065P.
PR	21-SEP-2000;	2000US-0234221P.
PR	21-SEP-2000;	2000US-0234222P.
PR	21-SEP-2000;	2000US-0234274P.
PR	23-SEP-2000;	2000US-02345997P.
PR	23-SEP-2000;	2000US-0234598P.
PR	26-SEP-2000;	2000US-0234584P.
PR	27-SEP-2000;	2000US-02335634P.
PR	27-SEP-2000;	2000US-02335635P.
PR	29-SEP-2000;	2000US-0233602P.
PR	02-OCT-2000;	2000US-02336802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241221P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	01-NOV-2000;	2000US-0244617P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246777P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246509P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	17-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.

PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250300P.
PR	01-DEC-2000;	2000US-0250316P.
PR	01-DEC-2000;	2000US-0250319P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251198P.
PR	05-DEC-2000;	2000US-0251988P.
PR	06-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0255479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2000US-0253678P.

Isolated polypeptide for treating, preventing and prognosing disorders related to the endocrine system including endocrine disorders, reproductive disorders, and gastrointestinal disorders and also for testing and detection e.g. diagnosis

Disclosure; SEQ ID NO 842; 558pp; English.

The invention relates to cDNAs encoding novel human endocrine antigens or a fragment having biological activity, a domain, an epitope, full length protein, variant, allelic variant or a species homologue of the cDNA/antigen. The DNAs and polypeptides are useful for preventing, treating or ameliorating a medical condition when administered (e.g. by gene therapy or antisense-therapy). Identifying mutations in the genes coding for the antigens is useful for diagnosing a pathological condition or a susceptibility to a pathological condition. The DNAs, antigens and antibodies raised against the antigens useful for treating, preventing and/or prognosing disorders related to the endocrine system or hormone imbalance or reproductive disorders, cancers of endocrine tissues, disorders of the pancreas (e.g. diabetes mellitus), the adrenal glands (e.g. hirsutism), ovaries, the thyroid (e.g. hyperthyroidism), the hypothalamus and testes (e.g. vanishing testes syndrome), many examples of diseases and disorders are given in the specification. The present sequence is genomic DNA fragment form a gene encoding an endocrine antigen of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at fp.wipo.int/pub/published pct sequences

Sequence 329 BP; 89 A; 56 C; 50 G; 134 T; 0 U; 0 Other;

```
Query Match      100.0%; Score 11; DB 4; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels
```

Qy 1 TAACTATATGA 11
Db 84 TAACTATATGA 74

RESULT 22
ADA30131
ID ADA30131 standard; DNA; 333 BP.

XX AC ADA30131;
 XX DT 20-NOV-2003 (first entry)
 XX DE DNA encoding Acinetobacter baumannii protein #1418.
 XX KW ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
 XX KW vaccine; plant biocontrol agent.
 XX OS Acinetobacter baumannii.
 XX PN US6562958-B1.
 XX PD 13-MAY-2003.
 XX PF 04-JUN-1999; 99US-00328352.
 XX PR 09-JUN-1998; 98US-0088701P.
 XX PA (GENO-) GENOME THERAPEUTICS CORP.
 XX PI Breton G, Bush D;
 XX DR WPI: 2003-576092/54.
 XX DR P-PSDB; ADA34257.
 XX PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 XX PS Example; SEQ ID NO 1418; 328pp; English.
 XX CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents DNA encoding an A. baumannii
 CC protein.
 XX CC
 XX SQ Sequence 333 BP; 117 A; 35 C; 79 G; 102 T; 0 U; 0 Other;
 Query Match 100.0%; Score 11; DB 9; Length 333;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TAACTATATGA 11
 Db 168 TAACTATATGA 178
 RESULT 23
 ADF02797
 ID ADF02797 standard; DNA; 336 BP.
 XX AC ADF02797;
 XX DT 12-FEB-2004 (first entry)
 XX DE Bacterial polynucleotide #3082.
 XX KW Proteus mirabilis infection; bacterial infection; antibacterial;
 XX KW immunostimulant; gene; ds.
 XX OS Proteus mirabilis.
 XX PN US6605709-B1.
 XX PD 12-AUG-2003.
 Query Match 100.0%; Score 11; DB 9; Length 333;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TAACTATATGA 11
 Db 168 TAACTATATGA 178
 RESULT 24
 ABV60195/C
 ID ABV60195 standard; cDNA; 346 BP.
 XX AC ABV60195;
 XX DT 13-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 60186.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US005171.
 XX PR 17-FEB-2000; 2000US-0183319P.
 XX PR 16-MAR-2000; 2000US-0189862P.
 XX PR 25-MAY-2000; 2000US-0207454P.
 XX PR 09-JUN-2000; 2000US-0211314P.
 XX PR 18-JUL-2000; 2000US-0219007P.
 XX PR 13-DEC-2000; 2000US-0255281P.
 XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PD

XX PF 05-APR-2000; 2000US-00543681.
 XX PR 09-APR-1999; 99US-0128706P.
 XX PA (GENO-) GENOME THERAPEUTICS CORP.
 XX PI Breton GL;
 XX DR WPI: 2003-895291/82.
 XX DR P-PSDB; ADF06969.
 XX PT New Proteus mirabilis polypeptides and polynucleotides, useful as
 PT reagents for diagnosis of bacterial disease, as components of
 PT antibacterial vaccines, as targets for antibacterial drugs, or as
 PT biocontrol agents for plants.
 XX PS Disclosure; SEQ ID NO 3082; 870pp; English.
 XX CC The invention relates to new Proteus mirabilis polypeptides and
 CC polynucleotides. The invention also relates to antibodies against the
 CC polypeptides, methods for producing the polypeptides, a method of
 CC generating vaccines for immunising an individual against P. mirabilis, a
 CC method for evaluating a compound for the ability to bind a P. mirabilis
 CC polypeptide and a method for screening test compounds for anti-bacterial
 CC activity. The polypeptides and polynucleotides are useful as molecular
 CC targets for diagnosing, preventing and treating pathological conditions
 CC resulting from bacterial infection, as reagents for diagnosis of
 CC bacterial diseases, as components of antibacterial vaccines, as targets
 CC for antibacterial drugs or as bio-control agents for plants. This
 CC sequence represents a Proteus mirabilis polynucleotide of the invention.
 XX CC
 XX SQ Sequence 336 BP; 129 A; 49 C; 61 G; 97 T; 0 U; 0 Other;
 Query Match 100.0%; Score 11; DB 10; Length 336;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TAACTATATGA 11
 Db 237 TAACTATATGA 247
 RESULT 24
 ABV60195/C
 ID ABV60195 standard; cDNA; 346 BP.
 XX AC ABV60195;
 XX DT 13-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 60186.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US005171.
 XX PR 17-FEB-2000; 2000US-0183319P.
 XX PR 16-MAR-2000; 2000US-0189862P.
 XX PR 25-MAY-2000; 2000US-0207454P.
 XX PR 09-JUN-2000; 2000US-0211314P.
 XX PR 18-JUL-2000; 2000US-0219007P.
 XX PR 13-DEC-2000; 2000US-0255281P.
 XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PD

PI Schlegel R, Endege WO, Monahan JB;
 DR WPI; 2001-662795/76.
 XX
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 XX Claim 1; Page 11480; 11750pp; English.
 XX
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABY00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 XX Sequence 346 BP; 101 A; 82 C; 60 G; 103 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 11; DB 5; Length 346;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAACTATATGA 11
 DB 52 TAACTATATGA 42
 |||||

RESULT 25
 AAS38386
 ID AAS38386 standard; cDNA; 347 BP.
 AC AAS38386;
 DT 17-DEC-2001 (first entry)
 XX Novel human diagnostic and therapeutic gene #1444.
 DE Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.
 XX Homo sapiens.
 OS
 PN WO200166753-A2.
 XX 13-SEP-2001.
 XX 09-MAR-2001; 2001WO-US007787.
 XX 09-MAR-2000; 2000US-0188609P.
 XX (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;
 PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
 PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
 XX WPI; 2001-530177/58.
 DR
 XX New polynucleotides and polypeptides, useful for diagnosis and treatment
 PT of breast, lung and colon cancer.
 PT
 XX Claim 1; Page 958; 1193pp; English.
 PS
 XX The invention relates to new polynucleotides and polypeptides, useful for
 CC diagnosis and treatment of breast, lung and colon cancer. The sequences

CC can be used in detecting differentially expressed genes correlated with a
 CC cancerous state of a mammalian cell, comprising detecting at least one
 CC differentially expressed gene product in a test sample derived from a
 CC cell suspected of being cancerous. They can also be used to inhibit
 CC tumour growth by modulating expression of a gene product. AAS36943-
 CC AAS39338 represent novel human diagnostic and therapeutic coding
 CC sequences of the invention
 XX
 SQ Sequence 347 BP; 124 A; 49 C; 53 G; 120 T; 0 U; 1 Other;
 Query Match 100.0%; Score 11; DB 4; Length 347;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAACTATATGA 11
 DB 104 TAACTATATGA 114
 |||||

RESULT 26
 ABL84127
 ID ABL84127 standard; cDNA; 348 BP.
 XX ABL84127;
 AC
 XX 17-MAY-2002 (first entry)
 DT Human ovarian cancer related cDNA clone SEQ ID NO:7105.
 DE Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
 XX Homo sapiens.
 OS
 PN WO200192581-A2.
 XX 06-DEC-2001.
 XX 29-MAY-2001; 2001WO-US017756.
 XX 26-MAY-2000; 2000US-0207484P.
 XX (CORI-) CORIXA CORP.
 PA Algate PA, Harlocker SL, Jones R;
 PI WPI; 2002-122075/16.
 DR
 XX Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PT polypeptide.
 XX Claim 1; SEQ ID NO 7105; 489pp; English.
 PS
 XX The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
 CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
 CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
 CC or antigen presenting cells that express (II). (I) has cytostatic
 CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
 CC detecting ovarian cancer in a patient's biological sample preferably
 CC serum or ovarian tissue. The method comprises contacting a biological
 CC sample from a patient with (IV), detecting the amount of polynucleotide
 CC hybridising to (IV) and comparing the amount to a predetermined cutoff
 CC value and thereby detecting ovarian cancer in the patient, where the
 CC amount of polynucleotide hybridising to (IV) is detected preferably by
 CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
 CC useful for stimulating and/or expanding T cells specific for an ovarian
 CC tumour protein comprising contacting T cells with (III) or (II). (III) is
 CC useful in design and preparation of ribozyme molecules for inhibiting
 CC expression of the tumour polypeptides and proteins in tumour cells; and
 CC to isolate a full length gene from a suitable library e.g., a tumour cDNA

CC library using well known techniques

XX Sequence 348 BP; 134 A; 46 C; 64 G; 104 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 11; DB 6; Length 348;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
 |||||
 Db 241 TAACTATATGCA 251

RESULT 27

ADFO0709 ID ADF00709 standard; DNA; 348 BP.
 XX AC ADF00709;
 XX DT 12-FEB-2004 (first entry)
 XX Bacterial polynucleotide #994.
 DE Proteus mirabilis infection; bacterial infection; antibacterial;
 KW immunostimulant; gene; ds.
 XX Proteus mirabilis.
 OS US6605709-B1.
 XX PN 12-AUG-2003.
 XX PF 05-APR-2000/ 2000US-00543681.
 XX PR 09-APR-1999/ 99US-0128706P.
 XX PA (GENO-) GENOME THERAPEUTICS CORP.
 XX PI Breton GL;

XX WPI; 2003-895291/82.
 DR P-PSDB; ADF04881.
 XX PT New Proteus mirabilis polypeptides and polynucleotides, useful as
 reagents for diagnosis of bacterial disease, as components of
 antibacterial vaccines, as targets for antibacterial drugs, or as
 biocontrol agents for plants.
 XX PS Disclosure; SEQ ID NO 994; 870pp; English.

XX The invention relates to new Proteus mirabilis polypeptides and
 polynucleotides. The invention also relates to antibodies against the
 polypeptides, methods for producing the polypeptides, a method of
 generating vaccines for immunising an individual against P. mirabilis, a
 method for evaluating a compound for the ability to bind a P. mirabilis
 polypeptide and a method for screening test compounds for anti-bacterial
 activity. The polypeptides and polynucleotides are useful as molecular
 targets for diagnosing, preventing and treating pathological conditions
 resulting from bacterial infection, as reagents for diagnosis of
 bacterial diseases, as components of antibacterial vaccines, as targets
 for antibacterial drugs or as bio-control agents for plants. This
 sequence represents a Proteus mirabilis polynucleotide of the invention.

XX Sequence 348 BP; 136 A; 61 C; 58 G; 93 T; 0 U; 0 Other;
 Query Match 100.0%; Score 11; DB 10; Length 348;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
 |||||
 Db 6 TAACTATATGCA 16

RESULT 28

ADH83614 ID ADH83614 standard; DNA; 351 BP.
 XX AC ADH83614;
 XX DT 22-APR-2004 (first entry)
 XX Enterococcus faecalis polynucleotide #1499.
 DE Enterococcus faecalis infection; transcription regulatory element;
 KW antibacterial; gene; ds.
 XX Enterococcus faecalis.
 OS US6617156-B1.
 XX PN 09-SEP-2003.
 XX PF 13-AUG-1998; 98US-00134000.
 XX PR 15-AUG-1997; 97US-0055778P.
 XX PA (DOUC/) DOUCETTE-STAMM L A.
 XX PI (BUSH/) BUSH D.

PI Doucette-Stamm LA, Bush D;

XX WPI; 2003-895394/82.

DR P-PSDB; ADH87019.

XX New nucleic acid comprising a sequence encoding an Enterococcus faecalis
 polypeptide, useful for preparing a composition for diagnosing or
 treating E. faecalis infection.

XX Disclosure; SEQ ID NO 1499; 193pp; English.

XX The invention relates to Enterococcus faecalis polynucleotides and
 polypeptides. The invention also relates to a recombinant expression
 vector comprising a polynucleotide operably linked to a transcription
 regulatory element, a cell comprising a recombinant vector, a method for
 producing an E. faecalis polypeptide, an isolated nucleic acid comprising
 a sequence not given in the specification, a recombinant vector
 comprising the nucleic acid and a cell comprising the recombinant vector.
 CC The polynucleotides can be used to detect the presence of E. faecalis in
 a sample. The sequences are useful for preparing a composition for
 diagnosing or treating Enterococcus faecalis infection. This sequence
 CC represents an E. faecalis polynucleotide of the invention.

XX Sequence 351 BP; 134 A; 60 C; 75 G; 82 T; 0 U; 0 Other;
 Query Match 100.0%; Score 11; DB 10; Length 351;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
 |||||
 Db 156 TAACTATATGCA 166

RESULT 29

AAK61214/c ID AAK61214 standard; cDNA; 360 BP.
 XX AC AAK61214;
 XX DT 06-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:6274.
 DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ss.

PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 PT WPI; 2001-483426/52.
 DR P-PSDB; AAM88433.
 DR
 XX
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 XX Claim 1; SEQ ID NO 6274; 3071pp + Sequence Listing; English.
 PS
 XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 XX
 XX
 SQ Sequence 360 BP; 106 A; 60 C; 42 G; 148 T; 0 U; 4 Other;
 Query Match 100.0%; Score 11; DB 4; Length 360;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAACTATATGCA 11
 DB 100 TAACTATATGCA 90
 |||||
 |||||
 RESULT 30
 ABV17250
 ID ABV17250 standard; cDNA; 370 BP.
 XX
 AC ABV17250;
 XX
 XX 13-SEP-2002 (first entry)
 DT
 XX
 DE Human prostate expression marker cDNA 17241.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 XX Homo sapiens.
 OS
 XX WO200160860-A2.
 PN
 XX 23-AUG-2001,
 PD
 XX 20-FEB-2001; 2001WO-US005171.
 PF
 XX 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA
 XX Schlegel R, Endege WO, Monahan JE;
 PI
 XX

DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 XX Claim 1; Page 2860; 11750pp; English.
 PS
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 XX Sequence 370 BP; 140 A; 64 C; 56 G; 110 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 11; DB 5; Length 370;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAACTATATGCA 11
 DB 124 TAACTATATGCA 134
 |||||
 |||||
 RESULT 31
 ACH30883/C
 ID ACH30883 standard; cDNA; 384 BP.
 XX
 AC ACH30883;
 XX
 DT 13-OCT-2003 (first entry)
 DT
 XX
 DE Human bone marrow cDNA #174.
 XX
 DE Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
 KW genome mapping; biodiversity; genetic disorder.
 KW
 XX Homo sapiens.
 OS
 XX US2003073623-A1.
 PN
 XX 17-APR-2003.
 PD
 XX 30-JUL-2001; 2001US-00918995.
 PF
 XX 30-JUL-2001; 2001US-00918995.
 PR
 XX (DMA/) DRMANAC R T.
 PA (LABA/) LABAT I.
 PA (STAC/) STACHE-CRAIN B.
 PA (DICK/) DICKSON M C.
 PA (JONE/) JONES L W.
 XX
 XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
 PI
 XX WPI; 2003-615964/58.
 DR
 XX New polynucleotide sequences obtained from various cDNA libraries, useful
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene
 PT mapping, in the recombinant production of protein, or in generating
 PT antisense DNA or RNA.
 PT
 XX Claim 1; SEQ ID NO 18095; 44pp; English.
 PS
 XX The invention relates to an isolated polynucleotide comprising any one of

CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SBH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide
 CC is useful for generating antibodies specific for it. The present sequence
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030073623
 XX
 XX SQ Sequence 384 BP; 130 A; 69 C; 55 G; 130 T; 0 U; 0 Other;

Query Match 100.0%; Score 11; DB 9; Length 384;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 60 TAACTATATGA 50

RESULT 32

ABQ57583/c
 ID ABQ57583 standard; cDNA; 392 BP.

XX AC ABQ57583;

XX DT 02-AUG-2002 (first entry)

XX DB Human colon cancer related nucleotide sequence SEQ ID NO:1278.

XX KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
 KW genetic analysis; diagnostic; antisense therapy; gene; ss.

XX OS Homo sapiens.

XX PN WO200229086-A2.

XX PD 11-APR-2002.

XX PF 02-OCT-2001; 2001WO-US030732.

XX PR 02-OCT-2000; 2000US-0237271P.

XX PA (FARB) BAYER CORP.

XX PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA,
 PI Thiagalingam A, Lewis ME;

XX DR WPI; 2002-426115/45.

XX PT New isolated nucleic acid that is differentially expressed in cancer
 PT tissue useful for determining the presence of colon cancer in a cell or
 PT tissue type, and in antisense therapy.

XX PS Claim 1; Fig 1; 796pp; English.

XX CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
 CC expressed in cancer tissues. ABQ7893 to ABQ7904 represent proteins
 CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
 CC used in antisense therapy. An antibody immunoreactive with a polypeptide
 CC encoded by (I) is useful for detecting cancer in a patient sample, and
 CC for detecting the presence or absence of a polynucleotide encoded by a
 CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
 CC from (I) can be used for determining the presence of a nucleic acid which
 CC hybridises to (I), and for determining the phenotype of cells in a sample

CC of cells from a patient. (I) is useful for determining the presence of
 CC colon cancer in a cell or tissue type, for determining the presence or
 CC state of other type of cancer, in antisense therapy, to generate
 CC macroarrays on a solid surface, to identify a chromosome on which the
 CC corresponding gene resides, and in tissue profiling, forensics, genetic
 CC analysis, mapping and diagnostic applications. (I) can be used to raise
 CC antibodies, and to screen for peptide analogues and antagonists
 XX
 XX SQ Sequence 392 BP; 129 A; 58 C; 43 G; 142 T; 0 U; 20 Other;

Query Match 100.0%; Score 11; DB 6; Length 392;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 209 TAACTATATGA 199

RESULT 33

AAH70968/c
 ID AAH70968 standard; cDNA; 395 BP.

XX AC AAH70968;

XX DT 19-SRP-2001 (first entry)

XX DB Human cervical cancer marker nucleic acid 2242.

XX KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX OS Homo sapiens.

XX PN WO200142467-A2.

XX PD 14-JUN-2001.

XX PF 08-DEC-2000; 2000WO-US033312.

XX PR 08-DEC-1999; 99US-0169681P.

XX PR 21-DEC-1999; 99US-0171350P.

XX PR 14-MAR-2000; 2000US-0189315P.

XX PR 12-MAY-2000; 2000US-0203791P.

XX PR 09-JUN-2000; 2000US-0210600P.

XX PR 21-JUL-2000; 2000US-0220114P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Deeds J, Berger A, Zhao X;

XX DR WPI; 2001-375006/39.

XX PS Claim 1; Page 474; 1051pp; English.

XX CC The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful: to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;
 CC and to select and/or assess the efficacy of a compound or therapy for
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy

XX SQ Sequence 395 BP; 101 A; 64 C; 77 G; 150 T; 0 U; 3 Other;

Query Match 100.0%; Score 11; DB 4; Length 395;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

Db 122 TAACTATATGA 112
|||||||
RESULT 34
AAK89394/c
ID AAK89394 standard; DNA; 397 BP.
XX AAK89394;
AC AAK89394;
XX
DT 05-NOV-2001 (first entry)
XX Human digestive system antigen genomic sequence SEQ ID NO: 2970.
DE Human; digestive system antigen; gene therapy; cancer; appendicitis;
XX ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ds.
XX Homo sapiens.
XX WO200155314-A2.
XX
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001324.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225475P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232403P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.

```

PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-502630/55.
XX
XX Polynucleotides encoding digestive system antigens, useful for
XX diagnosing, treating, preventing and/or prognosing disorders of the
XX digestive system, particularly cancer and cancer metastases.
XX
XX Disclosure; SEQ ID NO 2970; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human digestive system antigens. These can be used in the
XX diagnosis, treatment and prevention of digestive system disorders,
XX including cancer, Meckel's diverticulum, bacterial or parasitic
XX infections, appendicitis, Hirschsprung's disease, chronic colitis or
XX ulcerative colitis. The present sequence is a genomic DNA fragment
XX encoding a digestive system antigen of the invention
XX
XX SQ Sequence 397 BP; 114 A; 71 C; 111 G; 101 T; 0 U; 0 Other;
Query Match 100.0%; Score 11; DB 4; Length 397;
Best Local Similarity 100.0%; Pred. NO. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAACTATATGA 11
Db 227 TAACTATATGA 217
RESULT 35
AAK89949/c
ID AAK89949 standard; DNA; 397 BP.
XX
XX AAK89949;
AC
XX
XX 05-NOV-2001 (first entry)
XX
XX Human digestive system antigen genomic sequence SEQ ID NO: 3525.
XX
XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
XX ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
XX digestive system disorder; Meckel's diverticulum; ds.
XX
XX Homo sapiens.
XX
XX WO200155314-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001324.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 13-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216800P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0227009P.
PR 01-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232403P.
PR 14-SEP-2000; 2000US-0232404P.
PR 21-SEP-2000; 2000US-0232423P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.

```

PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249267P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251898P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-502630/55.

Polynucleotides encoding digestive system antigens, useful for
diagnosing, treating, preventing and/or prognosing disorders of the
digestive system, particularly cancer and cancer metastases.

Disclosure; SEQ ID NO 3525; 986pp; English.

CC The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a genomic DNA fragment
CC encoding a digestive system antigen of the invention
XX
SQ Sequence 397 BP; 114 A; 71 C; 111 G; 101 T; 0 U; 0 Other;
Query Match 100.0%; Score 11; DB 4; Length 397;
Best Local Similarity 100.0%; Pred No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAACTATATGA 11
Db 227 TAACTATATGA 217
RESULT 36
AAK88190/c
ID AAK88190 standard; cDNA; 397 BP.
XX
AC AAK88190;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen coding sequence SEQ ID NO: 506.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ss.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001324.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226689P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234233P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-502630/55.
XX P-PSDB; AAM92417.
XX
XX Polynucleotides encoding digestive system antigens, useful for
XX diagnosing, treating, preventing and/or prognosing disorders of the
XX digestive system, particularly cancer and cancer metastases.
XX
XX Claim 1; SEQ ID NO 506; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human digestive system antigens. These can be used in the
XX diagnosis, treatment and prevention of digestive system disorders,
XX including cancer, Meckel's diverticulum, bacterial or parasitic
XX infections, appendicitis, Hirschsprung's disease, chronic colitis or
XX ulcerative colitis. The present sequence is a cDNA encoding a digestive
XX system antigen of the invention
XX
XX Sequence 397 BP; 111 A; 71 C; 111 G; 102 T; 0 U; 2 Other;

Query Match 100.0%; Score 11; DB 4; Length 397;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
|||
Db 227 TAACTATATGA 217

RESULT 37
AAS31909/c
ID AAS31909 standard; DNA; 397 BP.
XX
XX AAS31909;
XX
XX 04-DEC-2001 (first entry)
XX
XX Human liver associated genomic DNA #83.

XX Liver associated protein; human; mouse; rabbit; goat; horse; cat; dog;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
anti-infertility.
XX
XX Homo sapiens.
XX
PN W0200155355-A1.
XX
XX
PD 02-AUG-2001;
XX
XX 17-JAN-2001; 2001WO-US001351.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226275P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251866P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX WPI; 2001-457728/49.
 XX
 XX Isolated nucleic acid molecule encoding a human liver related protein is
 PT used in preventing, treating or ameliorating disorders of the liver
 PT particularly cancer of the liver.
 XX
 XX Claim 1; SEQ ID NO 385; 526pp; English.
 XX
 XX Sequences AAS31827-AAS32182 represent genomic DNA molecules, which encode
 CC the liver associated polypeptides of the invention. Liver associated
 CC polypeptides and their associated polynucleotides are useful in the
 CC diagnosis, treatment and prevention of various types of disorders in e.g.
 CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A
 CC pathological condition can be determined by detecting the presence or
 CC absence of a mutation in a liver associated polynucleotide. The treatable
 CC disorders include autoimmune diseases such as rheumatoid arthritis,
 CC hyperproliferative disorders such as neoplasms of the breast or liver,
 CC cardiovascular disorders such as cardiac arrest, cerebrovascular
 CC disorders such as cerebral ischaemia, nervous system disorders such as
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
 CC ocular disorders such as corneal infection, endocrine disorders such as
 CC premature labour and infertility, gastrointestinal disorders such as
 CC Crohn's disease, renal disorders such as glomerulonephritis and
 CC respiratory disorders such as asthma and pleurisy. The polypeptides can
 CC also be used to aid wound healing, to prevent skin aging due to sunburn,
 CC to maintain organs before transplantation, to regenerate tissues and in
 CC chemotaxis. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 397 BP; 114 A; 71 C; 111 G; 101 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 11; DB 5; Length 397;
 Best Local Similarity 100.0%; Pred. NO. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TAACTATATCA 11
 Db 227 TAACTATATCA 217
 RESULT 38
 ABN90264/C
 ID ABN90264 standard; DNA; 397 BP.
 XX
 XX AC ABN90264;
 XX
 XX 24-JUL-2002 (first entry)
 DT
 XX Human liver antigen HLDOR73 genomic sequence, SEQ ID NO:385.
 DE
 XX Human; liver antigen; liver disorder; hepatic disorder; infection;
 KW hepatitis; viral; parasitic; bacterial; fungal; inflammatory condition;
 KW cirrhosis; granulomatous hepatitis; toxin damage; drug damage;
 KW autoimmune disease; Wilson's disease; primary biliary cirrhosis;
 KW neoplastic disorder; cancer; tumour; portal hypertension;
 KW gastrointestinal disorder; hepatitis; drug screening; gene therapy;
 KW

KW chromosome mapping; forensic analysis; antibody preparation;
 KW hepatotropic; cytostatic; antiinflammatory; virucide; antibacterial;
 KW fungicide; parasiticide; antidote; immunosuppressive; gene; ds.
 XX

OS Homo sapiens.

XX US2002042096-A1.

XX 11-APR-2002.

XX 17-JAN-2001; 2001US-00764887.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 28-JUN-2000; 2000US-0214886P.

PR 07-JUL-2000; 2000US-0216847P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 22-AUG-2000; 2000US-0226868P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 08-SEP-2000; 2000US-0231413P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 27-SEP-2000; 2000US-0235834P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 13-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241809P.

PR 01-NOV-2000; 2000US-0244617P.

PR 17-NOV-2000; 2000US-0249299P.

PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.

XX

XX (ROSE/) ROSEN C A.

PA (RUBE/) RUBEN S M.

PA (BARA/) BARASH S C.

XX

XX Rosen CA, Ruben SM, Barash SC;

XX WPI; 2002-381944/41.

XX New nucleic acid encoding human liver antigens, useful for diagnosis,

PT treatment and prevention of e.g. hepatitis and hepatic cancer, also

PT related polypeptides and antibodies.

XX Disclosure; SEQ ID NO 385; 181pp; English.

XX

XX The invention relates to 145 novel human liver antigens (ABP40831-ABP40975) and to cDNAs encoding them (ABN90036-ABN90180), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human liver antigen polynucleotides, antibodies against human liver antigens, and the use of liver antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various disorders of the liver. Such conditions include viral infections (e.g., cytomegalovirus, Epstein-Barr virus, hepatitis A virus, hepatitis B virus and hepatitis C virus), parasitic infections (e.g., Clonorchis sinensis, Echinococcus granulosus and Entamoeba histolytica), and also bacterial and fungal infections. Other disorders that may be treated include inflammatory conditions (e.g., cirrhosis and granulomatous hepatitis), damage caused by drugs or toxins, autoimmune diseases (e.g. Wilson's disease, primary biliary cirrhosis), neoplastic disorders (e.g., adenomas, haemangiomas and hepatocellular carcinoma), portal hypertension, or gastrointestinal disorders (e.g., peptic ulcers, gastritis and peritoneal diseases). Liver antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate liver antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as molecular weight markers or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. Sequences ABN90182-ABN90537 represent human liver antigen genomic sequences. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO at seqdata.uspto.gov/sequence/

XX SQ Sequence 397 BP; 114 A; 71 C; 111 G; 101 T; 0 U; 0 Other;

Query Match 100.0%; Score 11; DB 6; Length 397;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TAACTATATGCA 11
Db 227 TAACTATATGCA 217
|||||

RESULT 39
ABX86567/c
ID ABX86567 standard; cDNA; 397 BP.

XX AC ABX86567;

XX

XX 24-APR-2003 (first entry)

XX

DE Corn ear-derived polynucleotide (cpd) #5027.

XX

XX Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; structural gene; functional gene; regulatory gene;

XX corn ear-specific profile; gene transcription; gene expression;

XX hybrid plant; desirable trait expression; plant breeding program;

XX inheritance; desired characteristic; growth; development;

XX disease resistance; environmental adaptability; quality; yield;

XX multigene trait; plant; gene; ss.

XX

OS Zea mays.

XX

PN US6476212-B1.

XX

PN 05-NOV-2002.

XX

XX 14-MAY-1999; 99US-00313294.

XX

XX 26-MAY-1998; 98US-0086722P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

XX Lalgudi RV, Ito LY, Sherman BK;

XX WPI; 2003-208840/20.

XX

XX Novel purified corn-ear derived polynucleotide useful as hybridization probe for detecting polynucleotide in sample, and for identifying, growth, evaluating, and altering desired characteristics associated with growth, development.

XX Example; SEQ ID NO 5027; 390pp; English.

XX

XX The present invention relates to the isolation of corn ear-derived polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022 and SATMON023. Some of the cdps uniquely identify structural, functional, and regulatory genes of corn ear. The polynucleotide sequences are useful for detecting cdps in a sample, for producing a corn ear-specific profile of gene transcription, for detecting altered gene expression in inbred or hybrid plants, and for screening several molecules for specific binding to the polynucleotide. The cdps are useful to identify, isolate, or extend identical or related corn-ear nucleic acid sequences from DNA libraries, and in nucleic acid amplification or hybridisation techniques to follow the expression of desirable traits through plant breeding programs. Preferably, the cdps are used to identify, evaluate, alter, or follow the inheritance of desired characteristics associated with growth and development, disease resistance, environmental adaptability, quality, and yield of corn. The cdps are also useful as molecular markers for studying inheritance and multigene traits in a plant breeding program. The cdps are useful for producing purified corn-ear polypeptides by recombinant techniques. They are also useful in diagnostic assays to detect or confirm conditions or diseases associated with abnormal levels of cdp expression. ABX81541-ABX89140 represent corn ear-derived polynucleotides (cdps) of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site at seqdata.uspto.gov/psipsDIDEntry.html

XX SQ Sequence 397 BP; 88 A; 79 C; 86 G; 113 T; 0 U; 31 Other;

Query Match 100.0%; Score 11; DB 10; Length 397;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TAACTATATGCA 11
Db 132 TAACTATATGCA 122
|||||

RESULT 40
ADJ15177/c
ID ADJ15177 standard; DNA; 397 BP.

XX AC ADJ15177;

XX

XX 20-MAY-2004 (first entry)

XX

DE Human liver-related genomic DNA - SEQ ID 385.

XX

XX liver; virucide; fungicide; antibacterial; antiparasitic; hepatotropic; antiinflammatory; cytostatic; litholytic; antirheumatic; antiarthritic; neuroprotective; antidiabetic; anticoagulant; thrombolytic;

XX antiarteriosclerotic; cardiast; haemostatic; antiarrhythmic;

XX ophthalmological; antiarteriosclerotic; vasotropic; osteopathic;

XX nootropic; antiparkinsonian; anticonvulsant; neuroleptic; vasotropic;

XX cytosatic; gynaecological; viral; fungal; disease;

XX parasitic infection; cirrhosis; Wilson's disease;

XX gastrointestinal disorder; pancreatic; gallbladder; immune; blood;

XX hyperproliferative; cardiovascular; respiratory; musculoskeletal system;

XX neurological; endocrine; reproductive system; developmental; inherited;

XX human; ds.

XX

OS Homo sapiens.

XX

PN US2003077602-A1.
XX 24-APR-2003.
XX 14-FEB-2002; 2002US-00073961.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198113P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232337P.
PR 14-SEP-2000; 2000US-0232338P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241321P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0254097P.
PR 17-JAN-2001; 2001US-00764887.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX

PI Rosen CA, Ruben SM, Barash SC;
 DR WPI; 2003-765398/72.
 XX
 XX New liver related polypeptide, useful for diagnosis, treatment and/or
 PT prevention of liver, gastrointestinal, pancreatic, immune, blood related,
 PT endocrine, reproductive, hyperproliferative or reproductive disorders.
 XX
 PS Disclosure; SEQ ID NO 385; 181pp; English.
 XX
 XX The invention relates to a novel isolated, liver related polypeptide. The
 CC polypeptide of the invention demonstrates virucide, fungicide,
 CC antibacterial, antiparasitic, hepatotropic, antiinflammatory, cytostatic,
 CC litholytic, antirheumatic, antiarthritic, neuroprotective, antidiabetic,
 CC anticoagulant, thrombolytic, antiarteriosclerotic, cardiant, haemostatic,
 CC antiarrhythmic, ophthalmological, antiarteriosclerotic, vasotropic,
 CC osteopathic, nootropic, antiparkinsonian, anticonvulsant, neuroleptic,
 CC vasotropic, cytostatic and gynaecological activities. The polypeptides
 CC and polynucleotides of the invention may be useful for diagnosis,
 CC detection, treatment and/or prevention of disorders of the liver such as
 CC viral, fungal, bacterial or parasitic infections, cirrhosis, Wilson's
 CC disease, gastrointestinal disorders, pancreatic disorders, gallbladder
 CC diseases, immune disorders, blood related disorders, hyperproliferative
 CC disorders, cardiovascular disorders, respiratory disorders,
 CC musculoskeletal system disorders, neurological diseases, endocrine
 CC disorders, reproductive system disorders or developmental and inherited
 CC disorders. The current sequence is that of the human liver-related
 CC genomic DNA of the invention. The current sequence is not shown within
 CC the specification per se but was obtained electronically from the USPTO
 CC web-site.
 XX

Query Match 100.0%; Score 11; DB 11; Length 397;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAACTATATGCA 11
 |||||
 Db 227 TAACTATATGCA 217

RESULT 41
 ACN47251
 ID ACN47251 standard; cDNA; 399 BP.
 XX
 AC ACN47251;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 XX Cotton primed seed EST Clone ID: LIB3825-009-Q1-N6-A10, SEQ:2032.
 DE
 XX Cotton; plant; EST; expressed sequence tag; transgenic plant; seed;
 KW variety DP50B; library LIB3825; molecular tag; molecular marker;
 KW genetic mapping; molecular mapping; seed germination; plant growth;
 KW plant quality; plant yield; plant breeding; tissue printing; ss.
 XX
 OS Gossypium hirsutum.
 XX
 XX US2004123340-A1.
 PN
 XX 24-JUN-2004.
 XX
 XX 12-DEC-2001; 2001US-00021323.
 PF
 XX 14-DEC-2000; 2000US-0255619P.
 PR
 XX (DEIK/) DEIKMAN J.
 PA (FENG/) FENG P C C.
 PA (FINC/) FINCHER K L.
 PA (ZIEG/) ZIEGLER T E.
 XX
 XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;
 PI
 XX

DR WPI; 2004-479808/45.
 XX
 PT New isolated nucleic acid molecule that encodes a plant protein or its
 PT fragment, useful for isolating a variety of agronomically significant
 PT genes associated with plant growth, quality or yield, and as molecular
 PT tags to map genes.
 XX
 PS Claim 1; SEQ ID NO 2032; 34pp; English.
 XX
 XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
 CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
 CC from primed or non-primed seeds from variety DP50B, mature seeds from
 CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
 CC tissue, developing fibres, carpel walls and septa from variety
 CC Nucton338. The invention also relates to substantially purified
 CC proteins or their fragments encoded by nucleic acid molecules of the
 CC invention, and to transformed plants having a nucleic acid construct
 CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
 CC molecular tags to isolate genetic regions, to isolate genes, to map
 CC genes, to determine gene function and to determine whether genes are
 CC members of a particular gene family. The nucleic acid molecules may be
 CC used for isolating a variety of agronomically significant genes
 CC associated with plant growth, quality, yield, and could also serve as
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are
 CC also useful for identifying genes important in initiating and maintaining
 CC seed germination or that may be used to mitigate stresses encountered of
 CC during seed germination. The ESTs additionally enable the acquisition of
 CC promoters and cis-regulatory elements which will be useful to express
 CC agronomically significant genes in these tissues and/or other tissues,
 CC and also permits the acquisition of molecular markers useful in breeding
 CC schemes, genetic and molecular mapping, and in cloning of agronomically
 CC significant genes. The nucleic acid molecules are further useful for
 CC detecting the expression level or pattern of a protein or mRNA and for
 CC detecting the presence or quantity of a protein by tissue printing. The
 CC present sequence represents a specifically claimed EST isolated from a
 CC cotton variety DP50B primed seed cDNA library (LIB3825). The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from the US patent office at
 CC seqdata.uspto.gov/sequence.html?DocID=US20040123340
 XX
 SQ Sequence 399 BP; 130 A; 94 C; 71 G; 104 T; 0 U; 0 Other;
 Query Match 100.0%; Score 11; DB 13; Length 399;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAACTATATGCA 11
 |||||
 Db 255 TAACTATATGCA 265

RESULT 42
 AAI92400/C
 ID AAI92400 standard; cDNA; 400 BP.
 XX
 AC AAI92400;
 XX
 XX 06-NOV-2001 (first entry)
 DT
 XX Human polynucleotide SEQ ID NO 12460.
 DE
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200164835-A2.
 PN
 XX 07-SEP-2001.
 PD
 XX 26-FEB-2001; 2001WO-US004927.
 PF

XX 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
FI WPI; 2001-514838/56.
XX P-PSDB; AAO12469.
DR
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
PT
XX Claim 1; SEQ ID NO 12460; 1399pp + Sequence Listing; English.
PS
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 400 BP; 140 A; 79 C; 75 G; 106 T; 0 U; 0 Other;
Query Match 100.0%; Score 11; DB 4; Length 400;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAACTATATGA 11
Db 174 TAACTATATGA 164
RESULT 43
AAK59040/C
ID AAK59040 standard; cDNA; 403 BP.
XX
XX AAK59040;
XX
XX 06-NOV-2001 (first entry)
DT Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:4100.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytokinetic; gene therapy; vaccine; metastasis; ss.
XX Homo sapiens.
XX WO200157182-A2.
XX
XX 09-AUG-2001.
PD
XX 17-JAN-2001; 2001WO-US001354.
PF
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0234518P.
PR 14-AUG-2000; 2000US-0234519P.
PR 14-AUG-2000; 2000US-0235213P.
PR 14-AUG-2000; 2000US-0235214P.
PR 14-AUG-2000; 2000US-0235266P.
PR 14-AUG-2000; 2000US-0235267P.
PR 14-AUG-2000; 2000US-0235268P.
PR 14-AUG-2000; 2000US-0235270P.
PR 14-AUG-2000; 2000US-0235447P.
PR 14-AUG-2000; 2000US-0235757P.
PR 14-AUG-2000; 2000US-0235758P.
PR 14-AUG-2000; 2000US-0235759P.
PR 18-AUG-2000; 2000US-0236279P.
PR 22-AUG-2000; 2000US-0236581P.
PR 22-AUG-2000; 2000US-0236588P.
PR 22-AUG-2000; 2000US-0237182P.
PR 30-AUG-2000; 2000US-0237009P.
PR 01-SEP-2000; 2000US-0239287P.
PR 01-SEP-2000; 2000US-0239343P.
PR 01-SEP-2000; 2000US-0239344P.
PR 05-SEP-2000; 2000US-0239345P.
PR 05-SEP-2000; 2000US-0239509P.
PR 06-SEP-2000; 2000US-0239513P.
PR 06-SEP-2000; 2000US-0239437P.
PR 08-SEP-2000; 2000US-0239438P.
PR 08-SEP-2000; 2000US-0239422P.
PR 08-SEP-2000; 2000US-0239433P.
PR 08-SEP-2000; 2000US-0239434P.
PR 08-SEP-2000; 2000US-0239435P.
PR 12-SEP-2000; 2000US-0239681P.
PR 12-SEP-2000; 2000US-0239688P.
PR 14-SEP-2000; 2000US-0239397P.
PR 14-SEP-2000; 2000US-0239398P.
PR 14-SEP-2000; 2000US-0239399P.
PR 14-SEP-2000; 2000US-0239400P.
PR 14-SEP-2000; 2000US-0239401P.
PR 14-SEP-2000; 2000US-0239633P.
PR 14-SEP-2000; 2000US-0239634P.
PR 21-SEP-2000; 2000US-0239655P.
PR 21-SEP-2000; 2000US-0239656P.
PR 25-SEP-2000; 2000US-0239977P.
PR 25-SEP-2000; 2000US-0239988P.
PR 26-SEP-2000; 2000US-0239484P.
PR 27-SEP-2000; 2000US-0239583P.
PR 27-SEP-2000; 2000US-0239583P.
PR 29-SEP-2000; 2000US-0239627P.
PR 29-SEP-2000; 2000US-0239636P.
PR 29-SEP-2000; 2000US-0239636P.
PR 29-SEP-2000; 2000US-0239637P.
PR 02-OCT-2000; 2000US-0239680P.
PR 02-OCT-2000; 2000US-0239682P.
PR 02-OCT-2000; 2000US-0239703P.
PR 02-OCT-2000; 2000US-0239703P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251859P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 11-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX P-PSDB; AAM86259.
DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
DR useful for preventing, diagnosing and/or treating cancers and metastasis.
XX Claim 1; SEQ ID NO 4100; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to

CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 403 BP; 128 A; 53 C; 59 G; 155 T; 0 U; 8 Other;
Query Match 100.0%; Score 11; DB 4; Length 403;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAACTATATGA 11
Db 135 TAACTATATGA 125
RESULT 44
AAK74370/c
ID AAK74370 standard; DNA; 404 BP.
XX AAK74370;
AC
XX 07-NOV-2001 (first entry)
DT Human immune/hematopoietic antigen genomic sequence SEQ ID NO:29182.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
XX 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US001354.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.

```
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232081P.
PR 08-SEP-2000; 2000US-0232082P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232337P.
PR 14-SEP-2000; 2000US-0232338P.
PR 14-SEP-2000; 2000US-0232339P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234937P.
PR 25-SEP-2000; 2000US-0234938P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 29182; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention
XX
XX Sequence 404 BP; 140 A; 55 C; 49 G; 160 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 11; DB 4; Length 404;
XX Best Local Similarity 100.0%; Pred. No. 1.7e+03;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 TAACTATATGA 11
XX |||||
XX 132 TAACTATATGA 122
```

RESULT	45
ID	AAK78936
XX	AAK78936 standard; DNA; 415 BP.
AC	AAK78936;
XX	
DT	07-NOV-2001 (first entry)
XX	
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33748.
XX	
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW	Cytostatic; gene therapy; vaccine; metastasis; ds.
XX	
OS	Homo sapiens.
XX	
FN	WO200157182-A2.
XX	
PD	09-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US0011354.
XX	
PR	31-JAN-2000; 2000US-0179065P.
PR	04-FEB-2000; 2000US-0180628P.
PR	24-FEB-2000; 2000US-0184664P.
PR	02-MAR-2000; 2000US-0186350P.
PR	16-MAR-2000; 2000US-0189874P.
PR	17-MAR-2000; 2000US-0190076P.
PR	18-APR-2000; 2000US-0198123P.
PR	19-MAY-2000; 2000US-0205155P.
PR	07-JUN-2000; 2000US-0209467P.
PR	28-JUN-2000; 2000US-0214886P.
PR	30-JUN-2000; 2000US-0215135P.
PR	07-JUL-2000; 2000US-0216647P.
PR	07-JUL-2000; 2000US-0216880P.
PR	11-JUL-2000; 2000US-0217487P.
PR	11-JUL-2000; 2000US-0217496P.
PR	14-JUL-2000; 2000US-0218230P.
PR	26-JUL-2000; 2000US-0220963P.
PR	26-JUL-2000; 2000US-0220964P.
PR	14-AUG-2000; 2000US-0224518P.
PR	14-AUG-2000; 2000US-0224519P.
PR	14-AUG-2000; 2000US-0225213P.
PR	14-AUG-2000; 2000US-0225214P.
PR	14-AUG-2000; 2000US-0225266P.
PR	14-AUG-2000; 2000US-0225277P.
PR	14-AUG-2000; 2000US-0225268P.
PR	14-AUG-2000; 2000US-0225270P.
PR	14-AUG-2000; 2000US-0225477P.
PR	14-AUG-2000; 2000US-0225757P.
PR	14-AUG-2000; 2000US-0225758P.
PR	14-AUG-2000; 2000US-0225759P.
PR	18-AUG-2000; 2000US-0226279P.
PR	22-AUG-2000; 2000US-0226681P.
PR	22-AUG-2000; 2000US-0226688P.
PR	22-AUG-2000; 2000US-0227182P.
PR	23-AUG-2000; 2000US-0227009P.
PR	30-AUG-2000; 2000US-0228924P.
PR	01-SEP-2000; 2000US-0229287P.
PR	01-SEP-2000; 2000US-0229343P.
PR	01-SEP-2000; 2000US-0229344P.
PR	01-SEP-2000; 2000US-0229345P.
PR	03-SEP-2000; 2000US-0229509P.
PR	05-SEP-2000; 2000US-0229513P.
PR	06-SEP-2000; 2000US-0230437P.
PR	06-SEP-2000; 2000US-0230438P.
PR	08-SEP-2000; 2000US-0231242P.
PR	08-SEP-2000; 2000US-0231243P.
PR	08-SEP-2000; 2000US-0231413P.
PR	08-SEP-2000; 2000US-0231414P.
PR	08-SEP-2000; 2000US-0232080P.
PR	08-SEP-2000; 2000US-0232081P.
PR	13-SEP-2000; 2000US-0231968P.
PR	14-SEP-2000; 2000US-0233997P.
PR	14-SEP-2000; 2000US-0233998P.
PR	14-SEP-2000; 2000US-0233999P.
PR	14-SEP-2000; 2000US-0234000P.
PR	14-SEP-2000; 2000US-0234011P.
PR	14-SEP-2000; 2000US-0233063P.
PR	14-SEP-2000; 2000US-0233064P.
PR	14-SEP-2000; 2000US-0233065P.
PR	21-SEP-2000; 2000US-0234223P.
PR	21-SEP-2000; 2000US-0234224P.
PR	25-SEP-2000; 2000US-0234997P.
PR	25-SEP-2000; 2000US-0234998P.
PR	26-SEP-2000; 2000US-0235484P.
PR	27-SEP-2000; 2000US-0235834P.
PR	27-SEP-2000; 2000US-0235836P.
PR	29-SEP-2000; 2000US-0236327P.
PR	29-SEP-2000; 2000US-0236367P.
PR	29-SEP-2000; 2000US-0236368P.
PR	29-SEP-2000; 2000US-0236369P.
PR	29-SEP-2000; 2000US-0236370P.
PR	02-OCT-2000; 2000US-0236802P.
PR	02-OCT-2000; 2000US-0237037P.
PR	02-OCT-2000; 2000US-0237038P.
PR	02-OCT-2000; 2000US-0237039P.
PR	02-OCT-2000; 2000US-0237040P.
PR	13-OCT-2000; 2000US-0239935P.
PR	13-OCT-2000; 2000US-0239937P.
PR	20-OCT-2000; 2000US-0240960P.
PR	20-OCT-2000; 2000US-0241221P.
PR	20-OCT-2000; 2000US-0241785P.
PR	20-OCT-2000; 2000US-0241786P.
PR	20-OCT-2000; 2000US-0241787P.
PR	20-OCT-2000; 2000US-0241808P.
PR	20-OCT-2000; 2000US-0241809P.
PR	01-NOV-2000; 2000US-0244617P.
PR	08-NOV-2000; 2000US-0246474P.
PR	08-NOV-2000; 2000US-0246475P.
PR	08-NOV-2000;

PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Disclosure; SEQ ID NO 33748; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 XX
 SQ Sequence 415 BP; 133 A; 75 C; 116 G; 91 T; 0 U; 0 Other;
 Query Match 100.0%; Score 11; DB 4; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAACTATATGA 11
 DB 181 TAACTATATGA 191
 RESULT 46
 ABV49710
 ID ABV49710 standard; cDNA; 418 BP.
 XX
 AC ABV49710;
 XX
 DT 17-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 49701.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX

PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 9698; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 418 BP; 140 A; 89 C; 75 G; 114 T; 0 U; 0 Other;
 Query Match 100.0%; Score 11; DB 5; Length 418;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAACTATATGA 11
 DB 80 TAACTATATGA 90
 RESULT 47
 ABX46997/c
 ID ABX46997 standard; cDNA; 423 BP.
 XX
 AC ABX46997;
 XX
 DT 21-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #12162.
 XX
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX
 FN US2002137139-A1.
 XX
 PD 26-SEP-2002.
 XX
 PF 24-SEP-2001; 2001US-00960352.
 XX
 PR 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX
 PA (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX

PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX WPI; 2003-110599/10.
 XX
 XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX
 XX Claim 2; SEQ ID NO 12162; 245pp; English.
 XX
 XX The invention relates to a purified nucleic acid molecule associated with
 CC lactation and muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX
 XX Sequence 423 BP; 124 A; 61 C; 83 G; 155 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 11; DB 8; Length 423;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
 |||||
 Db 360 TAACTATATGCA 350

RESULT 48
 ABN73189
 ID ABN73189 standard; cDNA; 425 BP.
 XX
 XX ABN73189;
 AC
 XX
 XX 03-JUL-2002 (first entry)
 DT
 XX
 XX Bovine embryonic germ (EG) cell cDNA EST 000128a CONTIG 50.
 DE
 XX
 XX Bovine; Bos taurus; EST; expressed sequence tag; totipotence;
 KW development/ gene; ss.
 KW
 XX
 XX Bos taurus.
 OS
 XX
 XX WO200194550-A2.
 PN
 XX
 XX 13-DEC-2001.
 PD
 XX
 XX 07-JUN-2001; 2001WO-US018576.
 PF
 XX
 XX 07-JUN-2000; 2000US-0209874P.
 PR
 XX
 XX 06-JUN-2001; 2001US-00876143.
 PR
 XX
 XX (INFI-) INFINGEN INC.
 PA

XX
 PI Ellertsen KJ, Pfister-Genskow M, Childs L;
 XX WPI; 2002-351289/38.
 XX
 XX An expressed sequence tag (EST), the expression of which, or its
 PT complementary sequence, in a cell identifies the cell as a
 PT developmentally competent or incompetent cell.
 XX
 XX Example 16; Page 124; 584pp; English.
 PS
 XX
 XX The present invention describes an expressed sequence tag (EST), where
 CC the EST is an isolated, enriched, or purified nucleic acid sequence
 CC representing all or part of a gene, the expression of which, or its
 CC complementary sequence, in a cell identifies the cell as a
 CC developmentally competent or incompetent cell. Molecules which induce
 CC developmental competence in a cell line are useful for inducing
 CC totipotence in one or more cells. Molecules which induce developmental
 CC incompetence in a cell line are useful for preventing a full term
 CC pregnancy in an animal and inhibiting totipotence. The molecules are also
 CC useful for treating a disease in an animal by inducing development of one
 CC or more cells of the animal into a specific cell type. The present
 CC sequence represents a bovine EST which is given in the exemplification of
 CC the present invention
 XX
 XX Sequence 425 BP; 132 A; 81 C; 64 G; 140 T; 0 U; 8 Other;
 SQ

Query Match 100.0%; Score 11; DB 6; Length 425;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
 |||||
 Db 284 TAACTATATGCA 294

RESULT 49
 AAL25340/C
 ID AAL25340 standard; cDNA; 439 BP.
 XX
 XX AAL25340;
 AC
 XX
 XX 07-DEC-2001 (first entry)
 DT
 XX
 XX Human breast cancer expressed polynucleotide 17797.
 DE
 XX
 XX Human; breast cancer; cell marker; cytostatic; ss.
 KW
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200151628-A2.
 PN
 XX
 XX 19-JUL-2001.
 PD
 XX
 XX 10-JAN-2001; 2001WO-US000798.
 PF
 XX
 XX 14-JAN-2000; 2000US-0176077P.
 PR
 XX
 XX 14-MAR-2000; 2000US-0189167P.
 PR
 XX
 XX 24-MAR-2000; 2000US-0192099P.
 PR
 XX
 XX 29-MAR-2000; 2000US-0193480P.
 PR
 XX
 XX 15-MAY-2000; 2000US-0205230P.
 PR
 XX
 XX 09-JUN-2000; 2000US-0211315P.
 PR
 XX
 XX 25-JUL-2000; 2000US-0220534P.
 PR
 XX
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA
 XX
 XX Lillie J, Xu Y, Wang Y, Steinmann K;
 PI
 XX
 XX WPI; 2001-451856/48.
 DR
 XX
 XX New peptide useful as a marker for the diagnosis of breast cancer.
 PT
 XX
 XX Claim 1; Page 3295; 3695pp; English.
 PS

XX The invention relates to human breast cancer expressed polynucleotides
CC (AA107544-AA126789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
XX Sequence 439 BP; 147 A; 82 C; 67 G; 143 T; 0 U; 0 Other;
SQ Query Match 100.0%; Score 11; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAACTATATGA 11
Db 13 TAACTATATGA 3
RESULT 50
ABV47045
ID ABV47045 standard; cDNA; 439 BP.
XX AC ABV47045;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 47036.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
OS
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 9263; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 439 BP; 161 A; 84 C; 73 G; 121 T; 0 U; 0 Other;
SQ Query Match 100.0%; Score 11; DB 5; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAACTATATGA 11
Db 164 TAACTATATGA 174
RESULT 51
ADP95418/c
ID ADP95418 standard; cDNA; 440 BP.
XX AC ADP95418;
XX
DT 09-SEP-2004 (first entry)
XX
DE Cotton expressed sequence tag, EST, #4429.
XX
XX Cotton; ss; EST; expressed sequence tag; plant; plant protection;
KW plant improvement; marker-assisted breeding.
XX
XX Gossypium hirsutum; variety Nucleon33B.
OS
XX US2004123338-A1.
XX
XX 24-JUN-2004.
XX
XX 08-DEC-2000; 2000US-00732627.
XX
XX 10-DEC-1999; 99US-0170255P.
XX
XX (FINC/) FINCHER K L.
XX
XX Fincher KL;
XX
XX WPI; 2004-479807/45.
XX
XX New substantially purified nucleic acid molecule that encodes a cotton
PT protein or its fragment, useful as molecular tool for the targeting and
PT isolation of novel genes for plant protection and improvement.
XX
XX Claim 1; SEQ ID NO 4429; 30pp; English.
XX
XX The invention relates to a substantially purified nucleic acid molecule
CC that encodes a cotton protein or its fragment comprising an EST
CC (expressed sequence tag) appearing as ADP90990-ADP95919. Also included
CC are a substantially purified cotton protein or its fragment encoded by a
CC nucleic acid molecule above and a transformed plant (having a nucleic
CC acid molecule which comprises: an exogenous promoter region which
CC functions in a plant cell to cause the production of a mRNA molecule; a
CC structural nucleic acid molecule comprising one of the ESTs or their
CC complements; a 3' non-translated sequence that functions in the plant
CC cell to cause termination of transcription and addition of polyadenylated
CC ribonucleotides to a 3' end of the mRNA molecule). The ESTs are useful as
CC molecular tool for the targeting and isolation of novel genes for plant
CC protection and improvement. The ESTs are useful for developing new
CC strategies for understanding critical plant developmental and metabolic
CC pathways, for isolating genes and promoters, for identifying and mapping
CC the genes involved in developmental and metabolic pathways, and for
CC determining gene function. The cotton nucleic acid molecules are useful
CC as molecular tags to isolate genetic regions, isolate genes, map genes,
CC and determine gene function. The nucleic acid molecules are useful for
CC determining if genes are members of a particular gene family and for use
CC in marker-assisted breeding programs. The present sequence is one of the
CC 4930 cotton ESTs of the invention. NOTE: The sequences are not displayed
CC in the specification but are available in electronic format from the
CC USPTO at seqdata.uspto.gov/sequence.html?DocID=2004012338.
XX


```

PN WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 10163; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX SQ Sequence 464 BP; 154 A; 79 C; 81 G; 150 T; 0 U; 0 Other;
Query Match 100.0%; Score 11; DB 5; Length 464;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAACTATATGA 11
Db 245 TAACTATATGA 235
RESULT 55
AA110825/c
ID AA110825 standard; DNA; 465 BP.
XX AC AA110825;
XX 12-OCT-2001 (first entry)
XX DE Probe #758 for gene expression analysis in human cervical cell sample.
XX KW Probe; humah; microarray; gene expression; cervical epithelial cell;
XX human cervical cancer; ss.
XX OS Homo sapiens.
XX WO200157278-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000670.
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
XX
XX RESULT 56
ABAS2475/c
ID ABAS2475 standard; DNA; 465 BP.
XX AC ABAS2475;
XX 01-FEB-2002 (first entry)
XX DE Human foetal liver single exon nucleic acid probe #780.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX OS Homo sapiens.
XX WO200157277-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000669.
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234587P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for analyzing

```



```

QY      1 TAACTATATGA 11
DB      168 TAACTATATGA 158

RESULT 59
ABA22267/c
ID      ABA22267 standard; DNA; 465 BP.
XX
AC      ABA22267;
XX
DT      23-JAN-2002 (first entry)
XX
DE      Probe #733 for gene expression analysis in human heart cell sample.
XX
KW      Human; gene expression; heart; microarray; vascular system; probe;
KW      cardiovascular disease; hypertension; cardiac arrhythmia;
KW      congenital heart disease; ss.
XX
OS      Homo sapiens.
XX
PN      WO200157274-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US000666.
XX
PR      04-FEB-2000; 2000US-0180312P.
PR      26-MAY-2000; 2000US-0207456P.
PR      30-JUN-2000; 2000US-00608408.
PR      03-AUG-2000; 2000US-00632366.
PR      21-SEP-2000; 2000US-0234687P.
PR      27-SEP-2000; 2000US-0236359P.
PR      04-OCT-2000; 2000GB-00024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-488990/53.
XX
PT      Single exon nucleic acid probes for analyzing gene expression in human
PT      hearts.
XX
PS      Claim 1; SEQ ID NO 733; 530pp; English.
XX
CC      The present invention relates to single exon nucleic acid probes for
CC      measuring human gene expression in a sample derived from human heart. The
CC      present sequence is one such probe. The probes may be used for
CC      predicting, measuring and displaying gene expression in samples derived
CC      from the human heart via microarrays. By measuring gene expression, the
CC      probes are useful for predicting, diagnosing, grading, and staging,
CC      monitoring and prognosing diseases of the human heart and vascular system
CC      e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC      congenital heart disease. Note: The sequence data for this patent did not
CC      form part of the printed specification, but was obtained in electronic
CC      format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ      Sequence 465 BP; 134 A; 96 C; 74 G; 161 T; 0 U; 0 Other;

Query Match      100.0%; Score 11; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TAACTATATGA 11
DB      168 TAACTATATGA 158

RESULT 60
AAK26194/c
ID      AAK26194 standard; DNA; 465 BP.
XX
AC      AAK26194;
XX
DT      06-NOV-2001 (first entry)
XX
DE      Human bone marrow expressed single exon probe SEQ ID NO: 751.
XX
KW      Human; bone marrow expressed exon; gene expression analysis; probe;
KW      microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS      Homo sapiens.
XX
PN      WO200157276-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US000668.
XX
PR      04-FEB-2000; 2000US-0180312P.
PR      26-MAY-2000; 2000US-0207456P.
PR      30-JUN-2000; 2000US-00608408.
PR      03-AUG-2000; 2000US-00632366.
PR      21-SEP-2000; 2000US-0234687P.
PR      27-SEP-2000; 2000US-0236359P.
PR      04-OCT-2000; 2000GB-00024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WPI; 2001-488990/53.
XX
PT      Human genome-derived single exon nucleic acid probes useful for analyzing
PT      gene expression in human bone marrow.
XX
PS      Example 4; SEQ ID NO 751; 658pp + Sequence Listing; English.
XX
CC      The present invention provides a number of single exon nucleic acid
CC      probes which are derived from genomic sequences expressed in the human
CC      bone marrow. They can be used to measure gene expression in bone marrow
CC      samples, which may enable the improved diagnosis and treatment of cancers
CC      such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC      the probes of the invention
XX
SQ      Sequence 465 BP; 134 A; 96 C; 74 G; 161 T; 0 U; 0 Other;

Query Match      100.0%; Score 11; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TAACTATATGA 11
DB      168 TAACTATATGA 158

RESULT 61
AAK00742/c
ID      AAK00742 standard; DNA; 465 BP.
XX
AC      AAK00742;
XX
DT      05-NOV-2001 (first entry)
XX
DE      Human brain expressed single exon probe SEQ ID NO: 733.
XX
KW      Human; brain expressed exon; gene expression analysis; probe; microarray;
KW      Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW      ss.
XX
OS      Homo sapiens.
XX
PN      WO200157275-A2.
XX

```

```
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 733; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention
XX
XX Sequence 465 BP; 134 A; 96 C; 74 G; 161 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 11; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
DB 168 TAACTATATGCA 158

RESULT 62
ABS25784/C
ID ABS25784 standard; DNA; 465 BP.
XX
XX ABS25784;
AC
XX
XX 25-FEB-2003 (first entry)
DT
XX
XX Human liver single exon probe, SEQ ID No 774.
DE
XX
XX Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; BS.
KW
XX
XX Homo sapiens.
OS
XX
XX WO200157273-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000664.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
```

```
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-488998/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
XX Claim 1; SEQ ID NO 774; 658pp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABS25011-ABS51005 represent human
XX liver single exon nucleic acid probes of the invention. Note: The
XX sequence information for this patent does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 465 BP; 134 A; 96 C; 74 G; 161 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 11; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
DB 168 TAACTATATGCA 158

RESULT 63
AAI00750/C
ID AAI00750 standard; DNA; 465 BP.
XX
XX AAI00750;
AC
XX
XX 09-OCT-2001 (first entry)
DT
XX
XX Probe #741 used to measure gene expression in human breast sample.
DE
XX
XX Probe; human; breast disease; breast cancer; development disorder; BS;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
KW
XX
XX Homo sapiens.
OS
XX
XX WO200157270-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 29-JAN-2001; 2001WO-US000661.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-476286/51.
XX
XX Novel single exon nucleic acid probe used to measuring gene expression in
XX a human breast.
PT
```


XX PS Claim 25; SEQ ID NO 741; 322bp; English.

XX CC The present invention relates to novel single exon nucleic acid probes.

XX CC The present sequence is one such probe. The probes are useful for

CC measuring human gene expression in a human breast sample, where the probe

CC hybridises at high stringency to a nucleic acid expressed in the human

CC breast. The probes are useful for predicting, diagnosing, grading,

CC staging, monitoring and prognosing diseases of the human breast,

CC particularly those diseases with polygenic aetiology. The diseases

CC include: breast cancer, disorders of development, inflammatory diseases

CC of the breast, fibrocystic changes, proliferative breast disease and non-

CC carcinoma tumours. Note: The sequence data for this patent did not form

CC part of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 465 BP; 134 A; 96 C; 74 G; 161 T; 0 U; 0 Other;

Query Match 100.0%; Score 11; DB 5; Length 465;

Best Local Similarity 100.0%; Pred. No. 1.7e+03; Mismatches 0; Gaps 0;

Matches 11; Conservative 0; Indels 0;

Qy 1 TAACTATATGCA 11

Db 168 TAACTATATGCA 158

RESULT 64

ABS00780/C

ID ABS00780 standard; DNA; 465 BP.

XX AC ABS00780;

XX DT 19-AUG-2002 (first entry)

XX DE Human genome-derived single exon probe from lung SEQ ID NO 771.

XX KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

KW chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

KW hyaline membrane disease.

XX OS Homo sapiens.

XX PN WO200186003-A2.

XX PD 15-NOV-2001.

XX PF 30-JAN-2001; 2001WO-US000665.

XX PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX PF WPI; 2002-114183/15.

XX DR Spatially-addressable set of single exon nucleic acid probes, used to

PT measure gene expression in human lung samples.

XX PS Claim 1; SEQ ID NO 771; 634bp; English.

XX CC The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived

CC from human lung comprising single exon nucleic acid probes having one of

CC 12614 nucleic acid sequences mentioned in the specification, or their

CC complements or the 12387 open reading frames derived from the 12614

CC probes. Also included are a microarray comprising the novel set of probes

CC; the novel set of probes which hybridise at high stringency to a nucleic

CC acid expressed in the human lung; measuring gene expression in a sample

CC derived from human lung, comprising (a) contacting the array with a

CC collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of the

CC array; identifying exons in a eukaryotic genome, comprising (a)

CC algorithmically predicting at least one exon from genomic sequences of

CC the eukaryote; and (b) detecting specific hybridisation of detectably

CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

CC having a fragment identical to the predicted exon, the probe is included

CC in the above mentioned microarray; assigning exons to a single gene,

CC comprising (a) identifying exons from genomic sequence by the method

CC above and (b) measuring the expression of each of the exons in several

CC tissues and/or cell types using hybridisation to a single exon

CC microarrays having a probe with the exon, where a common pattern of

CC expression of the exons in the tissues and/or cell types indicates that

CC the exons should be assigned to a single gene; a peptide comprising one

CC of 12011 sequences, mentioned in the specification, or encoded by the

CC probes/open reading frames (ORF). The probes are used for gene expression

CC analysis, and for identifying exons in a gene, particularly using human

CC lung derived mRNA and for the study of lung diseases such as asthma, lung

CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung

CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,

CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-

CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary

CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,

CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary

CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The

CC present sequence is a single exon probe of the invention. Note: The

CC sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 465 BP; 134 A; 96 C; 74 G; 161 T; 0 U; 0 Other;

Query Match 100.0%; Score 11; DB 6; Length 465;

Best Local Similarity 100.0%; Pred. No. 1.7e+03; Mismatches 0; Gaps 0;

Matches 11; Conservative 0; Indels 0;

Qy 1 TAACTATATGCA 11

Db 168 TAACTATATGCA 158

RESULT 65

ACH24813

ID ACH24813 standard; cDNA; 470 BP.

XX AC ACH24813;

XX DT 13-OCT-2003 (first entry)

XX DE Human adult ovary cDNA #3193.

XX KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;

KW genome mapping; biodiversity; genetic disorder.

XX OS Homo sapiens.

XX PN US2003073623-A1.

XX PD 17-APR-2003.

XX PF 30-JUL-2001; 2001US-00918995.

XX PR 30-JUL-2001; 2001US-00918995.

```
XX (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
XX
PI Dmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 12025; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antisense DNA or RNA. The purified polypeptide
XX is useful for generating antibodies specific for it. The present sequence
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
XX Sequence 470 BP; 138 A; 85 C; 102 G; 136 T; 0 U; 9 Other;
XX
XX
XX Query Match 100.0%; Score 11; DB 9; Length 470;
XX Best Local Similarity 100.0%; Pred. No. 1.7e+03;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TAACTATATGCA 11
XX Db 304 TAACTATATGCA 314
XX
XX RESULT 66
XX ACH26788
XX ID ACH26788 standard; cDNA; 471 BP.
XX
XX AC ACH26788;
XX
XX 13-OCT-2003 (first entry)
XX
XX Human adult ovary cDNA #5168.
XX
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.
XX
XX Homo sapiens.
XX
XX US2003073623-A1.
XX
XX 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
XX (STAC/) STACHE-CRAIN B.
XX
XX (DRMA/) DICKSON M C.
XX (JONE/) JONES L W.
XX
XX Dmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 14000; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antisense DNA or RNA. The purified polypeptide
XX is useful for generating antibodies specific for it. The present sequence
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
XX Sequence 471 BP; 110 A; 109 C; 86 G; 161 T; 0 U; 5 Other;
XX
XX
XX Query Match 100.0%; Score 11; DB 9; Length 471;
XX Best Local Similarity 100.0%; Pred. No. 1.7e+03;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TAACTATATGCA 11
XX Db 264 TAACTATATGCA 274
XX
XX RESULT 67
XX AAI32986
XX ID AAI32986 standard; DNA; 476 BP.
XX
XX AC AAI32986;
XX
XX 17-OCT-2001 (first entry)
XX
XX Probe #1672 used to measure gene expression in human placenta sample.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632365.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
```

XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX Claim 25; SEQ ID NO 1672; 654pp; English.
XX The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
XX Sequence 476 BP; 120 A; 110 C; 106 G; 140 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 11; DB 4; Length 476;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAACTATATGA 11
DB 65 TAACTATATGA 75
RESULT 68
ABM42958
ID ABA42958 standard; DNA; 476 BP.
XX AC ABA42958;
XX 01-FEB-2002 (first entry)
XX Human breast cell single exon nucleic acid probe #1653.
XX Human; microarray; single exon probe; gene expression; breast; disease;
XX cancer; ss.
XX Homo sapiens.
XX WO200157271-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US0000662.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX Claim 1; SEQ ID NO 1653; 327pp + Sequence Listing; English.
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each

CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 476 BP; 120 A; 110 C; 106 G; 140 T; 0 U; 0 Other;
Query Match 100.0%; Score 11; DB 4; Length 476;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAACTATATGA 11
DB 65 TAACTATATGA 75
RESULT 69
ACH23189
ID ACH23189 standard; cDNA; 479 BP.
XX AC ACH23189;
XX 13-OCT-2003 (first entry)
XX Human adult ovary cDNA #1569.
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.
XX Homo sapiens.
XX US2003073623-A1.
XX 17-APR-2003.
XX 30-JUL-2001; 2001US-00918995.
XX 30-JUL-2001; 2001US-00918995.
XX (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
XX (STAC/) STACHE-CRAIN B.
XX (DICK/) DICKSON M C.
XX (JONE/) JONES L W.
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX Claim 1; SEQ ID NO 10401; 44pp; English.
XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversities, or in identifying mutations

CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030073623
XX
SQ Sequence 479 BP; 150 A; 82 C; 81 G; 163 T; 0 U; 3 Other;

Query Match 100.0%; Score 11; DB 9; Length 479;
Best Local Similarity 100.0%; Pred. NO. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
DB 360 TAACTATATGCA 370
|||||

RESULT 70
ABX60813/c
ID ABX60813 standard; DNA; 479 BP.

AC ABX60813;

XX
DT 26-FEB-2003 (first entry)

XX Arabidopsis thaliana polynucleotide #159.

XX Thale cress; gene; ds; genetic manipulation; plant; biosynthesis;
KW genetic modification; environmental stress; disease resistance;
KW fungicide; insecticide; stress tolerance.

XX Arabidopsis thaliana.

XX US2002142319-A1.

XX 03-OCT-2002.

XX 07-AUG-2001; 2001US-00924035.

XX 13-AUG-1999; 99US-0148784P.

XX 11-AUG-2000; 2000US-00638258.

XX (GORL/) GORLACH J.

XX (ANY/) AN Y.

XX (HAMI/) HAMILTON C M.

XX (PRIC/) PRICE J L.

XX (HARG/) HARGISS T R.

XX (YUY/) YU Y.

XX (RAME/) RAMAKA J G.

XX (PAGE/) PAGE A.

XX (MATH/) MATHAW A V.

XX (LED/) LEDFORD B L.

XX (WOES/) WOESSNER J P.

XX (HAAS/) HAAS W D.

XX (GARC/) GARCIA C A.

XX Gorlach J, An Y, Hamilton CM, Price JL, Hargiss TR, Yu Y;
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA;

XX WPI; 2003-102509/09.

XX Novel Arabidopsis thaliana nucleic acid useful for constructing a
PT transgenic plant with enhanced disease resistance and enhanced traits of
PT interest, as probes, and in diagnosis and screening purposes.

XX Claim 1; Page 59; 277pp; English.

XX The invention relates to Arabidopsis thaliana nucleic acid sequences. The

CC DNA sequences and the polypeptides they encode are useful for identifying
CC homologous or related genes, for producing compositions that modulate the
CC expression or function of the polypeptides, for mapping functional
CC regions of the protein, in diagnosis, for studying associated
CC physiological pathways, for genetic manipulation of cells, preferably
CC plant cells, in screening assays of various plant strains to determine
CC the strains that are capable of withstanding a particular disease or
CC environmental stress, for enhancing or inhibiting production of
CC biosynthetic products in plants and to create genetically modified and
CC transgenic organisms, such as plant cells and plants. Transgenic plants
CC are useful for introducing or improving disease resistance and stress
CC tolerance in plants, screening biologically active agents, such as
CC fungicides and insecticides, and for identifying factors involved in
CC biosynthetic pathways of nutritional, commercial or medicinal value.
CC Sequences ABX60653-ABX61554 represent Arabidopsis thaliana
CC polynucleotides of the invention
XX
SQ Sequence 479 BP; 148 A; 120 C; 80 G; 131 T; 0 U; 0 Other;

Query Match 100.0%; Score 11; DB 10; Length 479;
Best Local Similarity 100.0%; Pred. NO. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
DB 110 TAACTATATGCA 100
|||||

RESULT 71

ABV50770/c

ID ABV50770 standard; cDNA; 482 BP.

XX AC ABV50770;

XX 17-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 50761.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183119P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 9874; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX

SQL Sequence 482 BP; 126 A; 90 C; 70 G; 195 T; 0 U; 1 Other;

Query Match 100.0%; Score 11; DB 5; Length 482;

Best Local Similarity 100.0%; Pred. No. 1.7e+03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

Db 176 TAACTATATGA 166

RESULT 72

AAK89948/C

ID AAK89948 standard; DNA; 484 BP.

XX AAK89948;

AC AAK89948;

XX 05-NOV-2001 (first entry)

XX Human digestive system antigen genomic sequence SEQ ID NO: 3524.

XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
 KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
 KW digestive system disorder; Meckel's diverticulum; ds.

XX Homo sapiens.

XX WO200155314-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001324.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 14-AUG-2000; 2000US-0225758P.

XX 14-AUG-2000; 2000US-0225759P.

XX 18-AUG-2000; 2000US-0226279P.

XX 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239335P.
 PR 13-OCT-2000; 2000US-0239337P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249273P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250150P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-502630/55.
XX
XX Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
XX Disclosure; SEQ ID NO 3524; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a genomic DNA fragment
CC encoding a digestive system antigen of the invention
XX
SQ Sequence 484 BP; 149 A; 75 C; 122 G; 138 T; 0 U; 0 Other;

Query Match 100.0%; Score 11; DB 4; Length 484;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
Db 312 TAACTATATGCA 302
|||||

RESULT 73
ID AAK89391/C
ID AAK89391 standard; DNA; 484 BP.
XX
AC AAK89391;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen genomic sequence SEQ ID NO: 2967.
XX

KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
XX digestive system disorder; Meckel's diverticulum; db.
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001324.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.

```
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251859P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-502630/55.
XX Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX Disclosure; SEQ ID NO 2967; 986pp; English.
PS The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a genomic DNA fragment
CC encoding a digestive system antigen of the invention
XX
SQ Sequence 484 BP; 149 A; 75 C; 122 G; 138 T; 0 U; 0 Other;
Query Match 100.0%; Score 11; DB 4; Length 484;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAACTATATGA 11
Db 312 TAACTATATGA 302
RESULT 74
AAS31906/c
ID AAS31906 standard; DNA; 484 BP.
XX
AC AAS31906;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human liver associated genomic DNA #80.
XX
KW Liver associated protein; human; mouse; rabbit; goat; horse; cat; dog;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; nervous system disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
KW anti-infertility.
XX
OS Homo sapiens.
XX
PN WO200153355-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001351.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
```

PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218230P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236389P.
PR 29-SEP-2000; 2000US-0236390P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-457728/49.

Isolated nucleic acid molecule encoding a human liver related protein is used in preventing, treating or ameliorating disorders of the liver particularly cancer of the liver.

Claim 1; SEQ ID NO 382; 526pp; English.

Sequences AAS31827-AAS32182 represent genomic DNA molecules, which encode the liver associated polypeptides of the invention. Liver associated polypeptides and their associated polynucleotides are useful in the diagnosis, treatment and prevention of various types of disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A pathological condition can be determined by detecting the presence or

CC absence of a mutation in a liver associated polynucleotide. The treatable
 CC disorders include autoimmune diseases such as rheumatoid arthritis,
 CC hyperproliferative disorders such as neoplasms of the breast or liver,
 CC cardiovascular disorders such as cardiac arrest, cerebrovascular
 CC disorders such as cerebral ischaemia, nervous system disorders such as
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
 CC ocular disorders such as corneal infection, endocrine disorders such as
 CC premature labour and infertility, gastrointestinal disorders such as
 CC Crohn's disease, renal disorders such as glomerulonephritis and
 CC respiratory disorders such as asthma and pleurisy. The polypeptides can
 CC also be used to aid wound healing, to prevent skin aging due to sunburn,
 CC to maintain organs before transplantation, to regenerate tissues and in
 CC chemotaxis. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 484 BP; 149 A; 75 C; 122 G; 138 T; 0 U; 0 Other;

Query Match 100.0%; Score 11; DB 5; Length 484;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
 |||||
 Db 312 TAACTATATGA 302

RESULT 75

ABN90261/C
 ID ABN90261 standard; DNA; 484 BP.

AC ABN90261;

XX
 XX
 DT 24-JUL-2002 (first entry)

DE Human liver antigen HLDOR73 genomic sequence, SEQ ID NO:382.

XX Human; liver antigen; liver disorder; hepatic disorder; infection;
 KW hepatitis; viral; parasitic; bacterial; fungal; inflammatory condition;
 KW cirrhosis; granulomatous hepatitis; toxin damage; drug damage;
 KW autoimmune disease; Wilson's disease; primary biliary cirrhosis;
 KW neoplastic disorder; cancer; tumour; portal hypertension;
 KW gastrointestinal disorder; hepatitis; drug screening; gene therapy;
 KW chromosome mapping; forensic analysis; antibody preparation;
 KW hepatotropic; cytostatic; antiinflammatory; virucide; antibacterial;
 KW fungicide; parasiticide; antidote; immunosuppressive; gene; ds.

XX Homo sapiens.

XX US2002042096-A1.

XX 11-APR-2002.

XX 17-JAN-2001; 2001US-00764887.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 28-JUN-2000; 2000US-0214886P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 14-JUL-2000; 2000US-0217496P.

XX 26-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 14-AUG-2000; 2000US-0225758P.

PR 22-AUG-2000; 2000US-0226868P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.

(RUBE/) RUBEN S M.

(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2002-381944/41.

New nucleic acid encoding human liver antigens, useful for diagnosis, treatment and prevention of e.g. hepatitis and hepatic cancer, also related polypeptides and antibodies.

Disclosure; SEQ ID NO 382; 181pp; English.

The invention relates to 145 novel human liver antigens (ABP40831-ABP40975) and to cDNAs encoding them (ABN90036-ABN90180), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human liver antigen polynucleotides, antibodies against human liver antigens, and the use of liver antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various disorders of the liver. Such conditions include viral infections (e.g., cytomegalovirus, Epstein-Barr virus, hepatitis A virus, hepatitis B virus and hepatitis C virus), parasitic infections (e.g., Clonorchis sinensis, Echinococcus granulosus and Entamoeba histolytica), and also bacterial and fungal infections. Other disorders that may be treated include inflammatory conditions (e.g., cirrhosis and granulomatous hepatitis), damage caused by drugs or toxins, neoplastic diseases (e.g., Wilson's disease, primary biliary cirrhosis), autoimmune disorders (e.g., adenomas, haemangiomas and hepatocellular carcinoma), portal hypertension, or gastrointestinal disorders (e.g., peptic ulcers, gastritis and peritoneal diseases). Liver antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate liver antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as molecular weight markers or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. Sequences ABN90182-ABN90537 represent human liver antigen genomic sequences. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly
CC from the USPTO at seqdata.uspto.gov/sequence/

XX
SQ Sequence 484 BP; 149 A; 75 C; 122 G; 138 T; 0 U; 0 Other;
Query Match 100.0%; Score 11; DB 6; Length 484;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
| | | | | | | | | |
Db 312 TAACTATATGCA 302

Search completed: August 30, 2005, 05:58:47
Job time : 12.2234 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 05:29:20 ; Search time 8.25887 Seconds
(without alignments)
8714.755 Million cell updates/sec

Title: US-09-983-000a-1_COPY_1262_1272

Perfect score: 11

Sequence: 1 taactatatga ll

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 7331713 seqs, 3271544945 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10J_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	100.0	25	15	US-10-098-263B-83561 Sequence 83561, A
2	11	100.0	25	15	US-10-098-263B-112139 Sequence 112139, A
C 3	11	100.0	25	21	US-10-719-900-381915 Sequence 381915, A
C 4	11	100.0	25	21	US-10-719-900-425698 Sequence 425698, A
5	11	100.0	25	21	US-10-719-900-806378 Sequence 806378, A
C 6	11	100.0	25	21	US-10-956-157-211006 Sequence 211006, A
C 7	11	100.0	25	21	US-10-956-157-256232 Sequence 256232, A

25	11	100.0	25	22	US-10-719-956-280509 Sequence 280509, A
25	11	100.0	25	22	US-10-719-956-475932 Sequence 475932, A
25	11	100.0	25	22	US-10-719-956-475933 Sequence 475933, A
C 11	11	100.0	113	17	US-10-242-535A-2765 Sequence 2765, A
C 12	11	100.0	113	18	US-10-085-783A-2765 Sequence 2765, A
C 13	11	100.0	133	19	US-10-674-124A-11193 Sequence 11193, A
C 14	11	100.0	168	19	US-10-437-963-75263 Sequence 75263, A
C 15	11	100.0	175	19	US-10-674-124A-21027 Sequence 21027, A
C 16	11	100.0	180	9	US-09-969-373-93 Sequence 93, A
C 17	11	100.0	180	17	US-10-242-535A-4678 Sequence 4678, A
C 18	11	100.0	180	18	US-10-085-783A-4678 Sequence 4678, A
C 19	11	100.0	183	19	US-10-430-201-1741 Sequence 1741, A
C 20	11	100.0	183	19	US-10-430-201-1742 Sequence 1742, A
C 21	11	100.0	196	20	US-10-425-115-175786 Sequence 175786, A
C 22	11	100.0	201	19	US-10-741-601-21808 Sequence 21808, A
C 23	11	100.0	201	20	US-10-719-993-19827 Sequence 19827, A
C 24	11	100.0	201	20	US-10-719-993-23965 Sequence 23965, A
C 25	11	100.0	201	20	US-10-719-993-35222 Sequence 35222, A
C 26	11	100.0	201	21	US-10-719-993-47827 Sequence 47827, A
C 27	11	100.0	201	21	US-10-741-600-39662 Sequence 39662, A
C 28	11	100.0	201	21	US-10-741-600-52607 Sequence 52607, A
C 29	11	100.0	201	21	US-10-741-600-52760 Sequence 52760, A
C 30	11	100.0	201	21	US-10-741-600-61149 Sequence 61149, A
C 31	11	100.0	201	21	US-10-741-600-63869 Sequence 63869, A
C 32	11	100.0	217	9	US-09-983-965-3777 Sequence 3777, A
C 33	11	100.0	227	20	US-10-425-115-180092 Sequence 180092, A
C 34	11	100.0	231	18	US-10-424-599-23116 Sequence 23116, A
C 35	11	100.0	232	20	US-10-425-115-17351 Sequence 17351, A
C 36	11	100.0	235	19	US-10-021-323-1302 Sequence 1302, A
C 37	11	100.0	238	20	US-10-425-115-43347 Sequence 43347, A
C 38	11	100.0	239	9	US-09-923-876-3376 Sequence 3376, A
C 39	11	100.0	239	10	US-09-923-876-3376 Sequence 3376, A
C 40	11	100.0	240	17	US-10-242-535A-3405 Sequence 3405, A
C 41	11	100.0	240	18	US-10-085-783A-3405 Sequence 3405, A
C 42	11	100.0	265	17	US-10-242-535A-7185 Sequence 7185, A
C 43	11	100.0	265	18	US-10-085-783A-7185 Sequence 7185, A
C 44	11	100.0	268	10	US-09-754-853A-162 Sequence 162, A
C 45	11	100.0	269	9	US-09-969-373-772 Sequence 772, A
C 46	11	100.0	277	20	US-10-425-115-137251 Sequence 137251, A
C 47	11	100.0	288	20	US-10-425-115-58242 Sequence 58242, A
C 48	11	100.0	301	18	US-10-424-599-88082 Sequence 88082, A
C 49	11	100.0	307	17	US-10-242-535A-30437 Sequence 30437, A
C 50	11	100.0	307	18	US-10-085-783A-30437 Sequence 30437, A
C 51	11	100.0	313	18	US-10-424-599-66607 Sequence 66607, A
C 52	11	100.0	324	22	US-10-756-149-4480 Sequence 4480, A
C 53	11	100.0	326	20	US-10-425-115-9141 Sequence 9141, A
C 54	11	100.0	329	17	US-10-074-024-842 Sequence 842, A
C 55	11	100.0	330	20	US-10-425-115-155727 Sequence 155727, A
C 56	11	100.0	341	20	US-10-425-115-166717 Sequence 166717, A
C 57	11	100.0	346	20	US-10-357-930-60214 Sequence 60214, A
C 58	11	100.0	347	10	US-09-803-719-1444 Sequence 1444, A
C 59	11	100.0	348	9	US-09-867-701-7105 Sequence 7105, A
C 60	11	100.0	348	20	US-10-425-115-75052 Sequence 75052, A
C 61	11	100.0	349	18	US-10-424-599-41957 Sequence 41957, A
C 62	11	100.0	370	20	US-10-357-930-17241 Sequence 17241, A
C 63	11	100.0	382	19	US-10-437-963-74680 Sequence 74680, A
C 64	11	100.0	384	10	US-09-918-995-18095 Sequence 18095, A
C 65	11	100.0	384	18	US-10-424-599-82850 Sequence 82850, A
C 66	11	100.0	385	17	US-10-242-535A-14846 Sequence 14846, A
C 67	11	100.0	385	18	US-10-085-783A-14846 Sequence 14846, A
C 68	11	100.0	386	19	US-10-437-963-90941 Sequence 90941, A
C 69	11	100.0	387	17	US-10-242-535A-56366 Sequence 56366, A
C 70	11	100.0	387	18	US-10-085-783A-56366 Sequence 56366, A
C 71	11	100.0	392	11	US-09-969-034-1278 Sequence 1278, A
C 72	11	100.0	394	19	US-10-674-124A-3724 Sequence 3724, A
C 73	11	100.0	397	9	US-09-764-887-385 Sequence 385, A
C 74	11	100.0	397	14	US-10-073-961-385 Sequence 385, A
C 75	11	100.0	399	19	US-10-021-323-2032 Sequence 2032, A
C 76	11	100.0	400	17	US-10-242-535A-8865 Sequence 8865, A
C 77	11	100.0	400	18	US-10-085-783A-8865 Sequence 8865, A
C 78	11	100.0	400	20	US-10-425-115-28737 Sequence 28737, A
C 79	11	100.0	402	17	US-10-242-535A-33119 Sequence 33119, A
80	11	100.0	402	18	US-10-085-783A-33119 Sequence 33119, A

Sequence 108075, A
Sequence 97789, A
Sequence 20401, A
Sequence 49729, A
Sequence 39417, A
Sequence 70755, A
Sequence 39417, A
Sequence 50412, A
Sequence 12162, A
Sequence 139, App
Sequence 109098, A
Sequence 137546, A
Sequence 184006, A
Sequence 10468, A
Sequence 78495, A
Sequence 1679, App
Sequence 55531, A
Sequence 55531, A

ALIGNMENTS

RESULT 1
US-10-098-263B-83561
; Sequence 83561, Application US/10098263B
; Publication No. US2003010410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 83561
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-83561

Query Match 100.0%; Score 11; DB 15; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 7 TAACTATATGA 17

RESULT 2
US-10-098-263B-112139
; Sequence 112139, Application US/10098263B
; Publication No. US2003010410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 112139
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-112139

Query Match 100.0%; Score 11; DB 15; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 8 TAACTATATGA 18

RESULT 3
US-10-719-900-381915/c
; Sequence 381915, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 381915
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-381915

Query Match 100.0%; Score 11; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 17 TAACTATATGA 7

RESULT 4
US-10-719-900-425698/c
; Sequence 425698, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 425698
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-425698

Query Match 100.0%; Score 11; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 17 TAACTATATGA 7

RESULT 5
US-10-719-900-806378
; Sequence 806378, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:

; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 806378
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-806378

Query Match 100.0%; Score 11; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 4 TAACTATATGA 14

RESULT 6
US-10-956-157-211006/c
; Sequence 211006, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Mounts, William
; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 211006
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-211006

Query Match 100.0%; Score 11; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 24 TAACTATATGA 14

RESULT 7
US-10-956-157-256232/c
; Sequence 256232, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Mounts, William
; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 256232
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-256232

Query Match 100.0%; Score 11; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 14 TAACTATATGA 4

RESULT 8
US-10-719-956-280509/c
; Sequence 280509, Application US/10719956
; Publication No. US20040146910A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Rat
; FILE REFERENCE: 3527.1
; CURRENT APPLICATION NUMBER: US/10/719,956
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,836
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 699466
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 280509
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-719-956-280509

Query Match 100.0%; Score 11; DB 22; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 19 TAACTATATGA 9

RESULT 9
US-10-719-956-475932
; Sequence 475932, Application US/10719956
; Publication No. US20040146910A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Rat
; FILE REFERENCE: 3527.1
; CURRENT APPLICATION NUMBER: US/10/719,956
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,836
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 699466
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 475932
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-719-956-475932

Query Match 100.0%; Score 11; DB 22; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 2 TAACTATATGA 12

RESULT 10
US-10-719-956-475933
; Sequence 475933, Application US/10719956
; Publication No. US20040146910A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou

; TITLE OF INVENTION: Methods of Genetic Analysis of Rat
; FILE REFERENCE: 3527.1
; CURRENT APPLICATION NUMBER: US/10/719,956
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,836
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 699466
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 475933
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-719-956-475933

Query Match 100.0%; Score 11; DB 22; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
Db 2 TAACTATATGCA 12

RESULT 11
US-10-242-535A-2765/c
; Sequence 2765, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2765
; LENGTH: 113
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-2765

Query Match 100.0%; Score 11; DB 17; Length 113;
Best Local Similarity 100.0%; Pred. No. 5.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
Db 15 TAACTATATGCA 5

RESULT 12
US-10-085-783A-2765/c
; Sequence 2765, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017

; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2765
; LENGTH: 113
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-2765

Query Match 100.0%; Score 11; DB 18; Length 113;
Best Local Similarity 100.0%; Pred. No. 5.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
Db 15 TAACTATATGCA 5

RESULT 13
US-10-674-124A-11193
; Sequence 11193, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidetoshi
; APPLICANT: TAMIYA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-113699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 11193
; LENGTH: 133
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: DISD22_0004738
; FEATURE:
; OTHER INFORMATION: Located on chromosome 6
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; OTHER INFORMATION: sequence : 157101461
; FEATURE:
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 146256
US-10-674-124A-11193

Query Match 100.0%; Score 11; DB 19; Length 133;
Best Local Similarity 100.0%; Pred. No. 5.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
Db 81 TAACTATATGCA 91

RESULT 14
US-10-437-963-75263/c
; Sequence 75263, Application US/10437963

```
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 75263
; LENGTH: 168
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_75368C.1
US-10-437-963-75263

Query Match      100.0%; Score 11; DB 19; Length 168;
Best Local Similarity 100.0%; Pred. No. 5.1e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||
Db      164 TAACTATATGA 154

RESULT 15
US-10-674-124A-21027
; Sequence 21027, Application US/10674124A
; Publication No. US2004019797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidetoshi
; APPLICANT: TAMIYA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 21027
; LENGTH: 175
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 267617
; FEATURE:
; OTHER INFORMATION: Located on chromosome 14
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; OTHER INFORMATION: sequence : 88731454
; FEATURE:
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 882
US-10-674-124A-21027
```

```
Query Match      100.0%; Score 11; DB 19; Length 175;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||
Db      99 TAACTATATGA 109
```

```
RESULT 16
US-09-969-373-93
; Sequence 93, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
; APPLICANT: Effertz, Roger J.
; APPLICANT: Hauge, Brian M.
; TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
; FILE REFERENCE: 38-10(52679)A
; CURRENT APPLICATION NUMBER: US/09/969,373
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 09/754,853
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 09/760,427
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: US 09/855,768
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 93
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Glycine max
US-09-969-373-93
```

```
Query Match      100.0%; Score 11; DB 9; Length 180;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||
Db      45 TAACTATATGA 55
```

```
RESULT 17
US-10-242-535A-4678/c
; Sequence 4678, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4678
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-4678
```

```
Query Match      100.0%; Score 11; DB 17; Length 180;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TAACTATATGCA 11
Db      83 TAACTATATGCA 73

RESULT 18
US-10-085-783A-4678/c
; Sequence 4678, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 2001-07-13
; PRIOR FILING DATE: 2001-07-13
; PRIOR FILING DATE: 2001-03-12
; PRIOR FILING DATE: 2001-03-12
; PRIOR FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4678
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-4678

Query Match      100.0%; Score 11; DB 18; Length 180;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TAACTATATGCA 11
Db      83 TAACTATATGCA 73

RESULT 19
US-10-430-201-1741/c
; Sequence 1741, Application US/10430201
; Publication No. US20040162679A1
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; TITLE OF INVENTION: Method for Predicting Gene Potential and Cell Commitment
; FILE REFERENCE: 40716 (IP-010)
; CURRENT APPLICATION NUMBER: US/10/430,201
; CURRENT FILING DATE: 2003-05-05
; PRIOR FILING DATE: 2003-05-05
; PRIOR FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 4879
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1741
; LENGTH: 183
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-430-201-1741

Query Match      100.0%; Score 11; DB 19; Length 183;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TAACTATATGCA 11
Db      154 TAACTATATGCA 144

RESULT 20
US-10-430-201-1742/c
; Sequence 1742, Application US/10430201
; Publication No. US20040162679A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; TITLE OF INVENTION: Method for Predicting Gene Potential and Cell Commitment
; FILE REFERENCE: 40716 (IP-010)
; CURRENT APPLICATION NUMBER: US/10/430,201
; CURRENT FILING DATE: 2003-05-05
; PRIOR FILING DATE: 2003-05-05
; PRIOR FILING DATE: 2002-05-03
; NUMBER OF SEQ ID NOS: 4879
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1742
; LENGTH: 183
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-430-201-1742

Query Match      100.0%; Score 11; DB 19; Length 183;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TAACTATATGCA 11
Db      154 TAACTATATGCA 144

RESULT 21
US-10-425-115-175786
; Sequence 175786, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 175786
; LENGTH: 196
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MPT4577_91912C.1
US-10-425-115-175786

Query Match      100.0%; Score 11; DB 20; Length 196;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TAACTATATGCA 11
Db      146 TAACTATATGCA 156

RESULT 22
US-10-741-601-21808/c
; Sequence 21808, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21808
; LENGTH: 201
; TYPE: DNA
```



```
; ORGANISM: Homo sapiens
US-10-741-601-21808

Query Match      100.0%; Score 11; DB 19; Length 201;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      161 TAACTATATGA 151

RESULT 23
US-10-719-993-19827/c
; Sequence 19827, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19827
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-19827

Query Match      100.0%; Score 11; DB 20; Length 201;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      100 TAACTATATGA 90

RESULT 24
US-10-719-993-23965/c
; Sequence 23965, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23965
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-23965

Query Match      100.0%; Score 11; DB 20; Length 201;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      182 TAACTATATGA 172

RESULT 25
US-10-719-993-35222/c
; Sequence 35222, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 35222
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-35222

Query Match      100.0%; Score 11; DB 20; Length 201;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      161 TAACTATATGA 151

RESULT 26
US-10-719-993-47827/c
; Sequence 47827, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47827
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-47827

Query Match      100.0%; Score 11; DB 20; Length 201;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      62 TAACTATATGA 52

RESULT 27
US-10-741-600-39662/c
; Sequence 39662, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 39662
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-39662

Query Match      100.0%; Score 11; DB 21; Length 201;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 TAACTATATGA 11
|||||
Db 100 TAACTATATGA 90

RESULT 28

US-10-741-600-52607/c
; Sequence 52607, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52607
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-52607

Query Match 100.0%; Score 11; DB 21; Length 201;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
|||||
Db 40 TAACTATATGA 30

RESULT 29

US-10-741-600-52760/c
; Sequence 52760, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52760
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-52760

Query Match 100.0%; Score 11; DB 21; Length 201;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
|||||
Db 114 TAACTATATGA 104

RESULT 30

US-10-741-600-61149/c
; Sequence 61149, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22

; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61149
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-61149

Query Match 100.0%; Score 11; DB 21; Length 201;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
|||||
Db 161 TAACTATATGA 151

RESULT 31

US-10-741-600-63869/c
; Sequence 63869, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63869
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-63869

Query Match 100.0%; Score 11; DB 21; Length 201;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
|||||
Db 100 TAACTATATGA 90

RESULT 32

US-09-983-965-3777/c
; Sequence 3777, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 3777
; LENGTH: 217
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 46-LIB3058-036-Q1-K1-D6
US-09-983-965-3777

Query Match 100.0%; Score 11; DB 9; Length 217;

```
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 97 TAACTATATGA 87

RESULT 33
US-10-425-115-180092/c
; Sequence 180092, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 180092
; LENGTH: 227
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_95823C.1
US-10-425-115-180092

Query Match 100.0%; Score 11; DB 20; Length 227;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 117 TAACTATATGA 107

RESULT 34
US-10-424-599-23116
; Sequence 23116, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 23116
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_120876C.1
US-10-424-599-23116

Query Match 100.0%; Score 11; DB 18; Length 231;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 124 TAACTATATGA 134

RESULT 35
```

```
US-10-425-115-17351
; Sequence 17351, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 17351
; LENGTH: 232
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_115824C.1
US-10-425-115-17351

Query Match 100.0%; Score 11; DB 20; Length 232;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 208 TAACTATATGA 218

RESULT 36
US-10-021-323-1902
; Sequence 1902, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 1902
; LENGTH: 235
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3825-007-Q1-N6-C8
US-10-021-323-1902

Query Match 100.0%; Score 11; DB 19; Length 235;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 95 TAACTATATGA 105

RESULT 37
US-10-425-115-43347
; Sequence 43347, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 43347
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_13953C.1
US-10-425-115-43347

Query Match 100.0%; Score 11; DB 20; Length 238;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
| | | | | | | | | |
Db 61 TAACTATATGA 71

RESULT 38
US-09-923-876-3376
; Sequence 3376, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 3376
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700162152H1
US-09-923-876-3376

Query Match 100.0%; Score 11; DB 9; Length 239;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
| | | | | | | | | |
Db 29 TAACTATATGA 39

RESULT 39
US-09-923-876-3376
; Sequence 3376, Application US/09923876
; Publication No. US20030237110A9
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329

; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 3376
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. US20030237110A9 700162152H1
US-09-923-876-3376

Query Match 100.0%; Score 11; DB 10; Length 239;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
| | | | | | | | | |
Db 29 TAACTATATGA 39

RESULT 40
US-10-242-535A-3405/c
; Sequence 3405, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3405
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-3405

Query Match 100.0%; Score 11; DB 17; Length 240;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
| | | | | | | | | |
Db 161 TAACTATATGA 151

RESULT 41
US-10-085-783A-3405/c
; Sequence 3405, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017

; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3405
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-3405

Query Match 100.0%; Score 11; DB 18; Length 240;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||
Db 161 TAACTATATGA 151

RESULT 42

US-10-242-535A-7185/c
; Sequence 7185, Application US/10242535A
; Publication No. US20040013663A1

; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.

; APPLICANT: Liew, C.C.

; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis

; FILE REFERENCE: 4231/2005

; CURRENT APPLICATION NUMBER: US/10/242,535A

; CURRENT FILING DATE: 2002-09-12

; PRIOR APPLICATION NUMBER: US 10/085,783

; PRIOR FILING DATE: 2002-02-28

; PRIOR APPLICATION NUMBER: US 60/305,340

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: US 60/275,017

; PRIOR FILING DATE: 2001-03-12

; PRIOR APPLICATION NUMBER: US 60/271,955

; PRIOR FILING DATE: 2001-02-28

; NUMBER OF SEQ ID NOS: 58994

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 7185

; LENGTH: 265

; TYPE: DNA

; ORGANISM: Human

US-10-242-535A-7185

Query Match 100.0%; Score 11; DB 17; Length 265;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||
Db 199 TAACTATATGA 189

RESULT 43

US-10-085-783A-7185/c
; Sequence 7185, Application US/10085783A
; Publication No. US20040037841A1

; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.

; APPLICANT: Liew, C.C.

; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis

; FILE REFERENCE: 4231/2002

; CURRENT APPLICATION NUMBER: US/10/085,783A

; CURRENT FILING DATE: 2002-02-28

; PRIOR APPLICATION NUMBER: US 60/305,340

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: US 60/275,017

; PRIOR FILING DATE: 2001-03-12

; PRIOR APPLICATION NUMBER: US 60/271,955

; PRIOR FILING DATE: 2001-02-28

; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7185
; LENGTH: 265
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-7185

Query Match 100.0%; Score 11; DB 18; Length 265;
Best Local Similarity 100.0%; Pred. No. 5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||
Db 199 TAACTATATGA 189

RESULT 44

US-09-754-853A-162

; Sequence 162, Application US/09754853A

; Publication No. US20030005491A1

; GENERAL INFORMATION:

; APPLICANT: Hauge, Brian M.

; APPLICANT: Parnell, Laurence D.

; APPLICANT: Parsons, Jeremy D.

; APPLICANT: Wang, Ming Li

; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With

; FILE REFERENCE: 38-10(15810)B

; CURRENT APPLICATION NUMBER: US/09/754,853A

; CURRENT FILING DATE: 2001-01-05

; PRIOR APPLICATION NUMBER: US 60/174,880

; PRIOR FILING DATE: 2000-01-07

; NUMBER OF SEQ ID NOS: 1119

; SEQ ID NO 162

; LENGTH: 268

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: 240017_region_G3_4566_16

US-09-754-853A-162

Query Match 100.0%; Score 11; DB 10; Length 268;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||
Db 40 TAACTATATGA 50

RESULT 45

US-09-969-373-772

; Sequence 772, Application US/09969373

; Patent No. US20020133852A1

; GENERAL INFORMATION:

; APPLICANT: Effertz, Roger J.

; APPLICANT: Hauge, Brian M.

; TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping

; FILE REFERENCE: 38-10(52679)A

; CURRENT APPLICATION NUMBER: US/09/969,373

; CURRENT FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US 09/754,853

; PRIOR FILING DATE: 2001-01-05

; PRIOR APPLICATION NUMBER: US 09/760,427

; PRIOR FILING DATE: 2001-01-13

; PRIOR APPLICATION NUMBER: US 09/855,768

; NUMBER OF SEQ ID NOS: 4593

; SEQ ID NO 772

; LENGTH: 269

; TYPE: DNA

; ORGANISM: Glycine max

US-09-969-373-772

Query Match 100.0%; Score 11; DB 9; Length 269;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
|||||
Db 42 TAACTATATGA 52

RESULT 46

US-10-425-115-137251
; Sequence 137251, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 137251
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_56650C.1
US-10-425-115-137251

Query Match 100.0%; Score 11; DB 20; Length 277;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
|||||
Db 239 TAACTATATGA 249

RESULT 47

US-10-425-115-584242/c
; Sequence 584242, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 58242
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_153113C.1
US-10-425-115-584242

Query Match 100.0%; Score 11; DB 20; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
|||||
Db 174 TAACTATATGA 164

RESULT 48

US-10-424-599-88082
; Sequence 88082, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 88082
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_50546C.1
US-10-424-599-88082

Query Match 100.0%; Score 11; DB 18; Length 301;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
|||||
Db 117 TAACTATATGA 127

RESULT 49

US-10-242-535A-30437
; Sequence 30437, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30437
; LENGTH: 307
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4)-(4)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-30437

Query Match 100.0%; Score 11; DB 17; Length 307;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
|||||
Db 115 TAACTATATGA 125

```
RESULT 50
US-10-085-783A-30437
; Sequence 30437, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/271,955
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30437
; LENGTH: 307
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-30437

Query Match 100.0%; Score 11; DB 18; Length 307;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
Db 115 TAACTATATGCA 125

RESULT 51
US-10-424-599-66607
; Sequence 66607, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 66607
; LENGTH: 313
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_3115C.1
US-10-424-599-66607

Query Match 100.0%; Score 11; DB 18; Length 313;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
Db 238 TAACTATATGCA 248

RESULT 52
US-10-756-149-4480
```

```
; Sequence 4480, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4480
; LENGTH: 324
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-4480

Query Match 100.0%; Score 11; DB 22; Length 324;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
Db 273 TAACTATATGCA 283

RESULT 53
US-10-425-115-9141
; Sequence 9141, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 9141
; LENGTH: 326
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_108330C.1
US-10-425-115-9141

Query Match 100.0%; Score 11; DB 20; Length 326;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
Db 202 TAACTATATGCA 212

RESULT 54
US-10-074-024-842/c
; Sequence 842, Application US/10074024
; Publication No. US20030232975A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC001C1
; CURRENT APPLICATION NUMBER: US/10/074,024
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 879
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 842
```

```
; LENGTH: 329
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-074-024-842

Query Match      100.0%; Score 11; DB 17; Length 329;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 84 TAACTATATGA 74

RESULT 55
US-10-425-115-155727/c
; Sequence 155727, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 155727
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_73608C.1
US-10-425-115-155727

Query Match      100.0%; Score 11; DB 20; Length 330;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 212 TAACTATATGA 202

RESULT 56
US-10-425-115-166717/c
; Sequence 166717, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 166717
; LENGTH: 341
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_83624C.1
US-10-425-115-166717

Query Match      100.0%; Score 11; DB 20; Length 341;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 TAACTATATGA 11
Db 261 TAACTATATGA 251

RESULT 57
US-10-357-930-60214/c
; Sequence 60214, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60214
; LENGTH: 346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-60214

Query Match      100.0%; Score 11; DB 20; Length 346;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 52 TAACTATATGA 42

RESULT 58
US-09-803-719-1444
; Sequence 1444, Application US/09803719
; Publication No. US20030044783A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
```


; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Human Genes and Gene Products
; FILE REFERENCE: 1624.002
; CURRENT APPLICATION NUMBER: US/09/803,719
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/188,609
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 2396
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1444
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(347)
; OTHER INFORMATION: n = A,T,C or G
US-09-803-719-1444

Query Match 100.0%; Score 11; DB 10; Length 347;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 104 TAACTATATGA 114

RESULT 59

US-09-867-701-7105
; Sequence 7105, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7105
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-7105

Query Match 100.0%; Score 11; DB 9; Length 348;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 241 TAACTATATGA 251

RESULT 60

US-10-425-115-75052/c
; Sequence 75052, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 75052
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_168469C.1
US-10-425-115-75052

Query Match 100.0%; Score 11; DB 20; Length 348;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 263 TAACTATATGA 253

RESULT 61

US-10-424-599-41957
; Sequence 41957, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 41957
; LENGTH: 349
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_13788C.1
US-10-424-599-41957

Query Match 100.0%; Score 11; DB 18; Length 349;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 252 TAACTATATGA 262

RESULT 62

US-10-357-930-17241
; Sequence 17241, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454

; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17241
; LENGTH: 370
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-17241

Query Match 100.0%; Score 11; DB 20; Length 370;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 124 TAACTATATGA 134

RESULT 63

US-10-437-963-74680
; Sequence 74680, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 74680
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_74842C.1
US-10-437-963-74680

Query Match 100.0%; Score 11; DB 19; Length 382;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 334 TAACTATATGA 344

RESULT 64

US-09-918-995-18095/c
; Sequence 18095, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18095
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-18095

Query Match 100.0%; Score 11; DB 10; Length 384;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 60 TAACTATATGA 50

RESULT 65

US-10-424-599-82850
; Sequence 82850, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 82850
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_45831C.1
US-10-424-599-82850

Query Match 100.0%; Score 11; DB 18; Length 384;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 255 TAACTATATGA 265

RESULT 66

US-10-242-535A-14846
; Sequence 14846, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14846
; LENGTH: 385
; TYPE: DNA

```
; ORGANISM: Human
US-10-242-535A-14846

Query Match      100.0%; Score 11; DB 17; Length 385;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      264 TAACTATATGA 274

RESULT 67
US-10-085-783A-14846
; Sequence 14846, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 14846
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Human

Query Match      100.0%; Score 11; DB 18; Length 385;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      264 TAACTATATGA 274

RESULT 68
US-10-437-963-90941
; Sequence 90941, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)/B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 90941
; LENGTH: 386
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_89564C.1
US-10-437-963-90941
```

```
Query Match      100.0%; Score 11; DB 19; Length 386;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      105 TAACTATATGA 115

RESULT 69
US-10-242-535A-56366/c
; Sequence 56366, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 56366
; LENGTH: 387
; TYPE: DNA
; ORGANISM: Human

Query Match      100.0%; Score 11; DB 17; Length 387;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      91 TAACTATATGA 81

RESULT 70
US-10-085-783A-56366/c
; Sequence 56366, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 56366
```


GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P113C1
; CURRENT APPLICATION NUMBER: US/10/073,961
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/764,887
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17

; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match 100.0%; Score 11; DB 14; Length 397;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||
Db 227 TAACTATATGA 217

RESULT 75
US-10-021-323-2032
; Sequence 2032, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 2032
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3825-009-Q1-N6-A10
US-10-021-323-2032

Query Match 100.0%; Score 11; DB 19; Length 399;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||
Db 255 TAACTATATGA 265

Search completed: August 30, 2005, 13:38:58
Job time : 8.25887 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 04:09:37 ; Search time 1.93262 Seconds
(without alignments)
9313.267 Million cell updates/sec

Title: US-09-983-000A-1_COPY_1262_1272

Perfect score: 11

Sequence: 1 taactatatga 11

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents NA.*

- 1: /cgm2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgm2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgm2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgm2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgm2_6/ptodata/1/ina/PTCUS COMB.seq.*
- 6: /cgm2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	100.0	158	4	US-09-513-999C-15732
2	11	100.0	220	4	US-09-513-999C-32605
3	11	100.0	255	4	US-09-513-999C-12608
4	11	100.0	284	4	US-09-270-767-31378
5	11	100.0	333	4	US-09-328-352-1418
6	11	100.0	336	4	US-09-543-681A-3082
7	11	100.0	348	4	US-09-543-681A-994
8	11	100.0	351	4	US-09-134-000C-1499
9	11	100.0	397	4	US-09-313-294A-5027
10	11	100.0	422	4	US-09-621-976-3695
11	11	100.0	460	4	US-09-513-999C-13380
12	11	100.0	542	4	US-09-270-767-7858
13	11	100.0	542	4	US-09-270-767-23140
14	11	100.0	561	1	US-08-480-784-29
15	11	100.0	561	1	US-08-483-553-29
16	11	100.0	561	1	US-08-487-002-29
17	11	100.0	561	1	US-08-483-554B-29
18	11	100.0	561	1	US-08-488-011B-29
19	11	100.0	561	3	US-08-850-727-29
20	11	100.0	561	5	PCT-US95-10202-29
21	11	100.0	561	5	PCT-US95-10203-29
22	11	100.0	561	5	PCT-US95-10220-29
23	11	100.0	570	4	US-09-328-352-3569
24	11	100.0	578	3	US-09-020-356-104
25	11	100.0	578	3	US-09-030-607-104
26	11	100.0	578	3	US-09-439-313-104
27	11	100.0	578	3	US-09-352-616A-104

3	578	11	100.0	US-09-232-149A-104	Sequence 104, App
4	578	11	100.0	US-09-159-812-104	Sequence 104, App
5	578	11	100.0	US-09-636-215-104	Sequence 104, App
6	578	11	100.0	US-09-585-166A-104	Sequence 104, App
7	578	11	100.0	US-09-115-453-104	Sequence 104, App
8	578	11	100.0	US-09-688-489-104	Sequence 104, App
9	578	11	100.0	US-09-679-426-104	Sequence 104, App
10	578	11	100.0	US-09-759-143-104	Sequence 104, App
11	578	11	100.0	US-09-651-236-104	Sequence 104, App
12	578	11	100.0	US-09-030-607-204	Sequence 204, App
13	578	11	100.0	US-09-439-313-204	Sequence 204, App
14	578	11	100.0	US-09-352-616A-204	Sequence 204, App
15	578	11	100.0	US-09-232-149A-204	Sequence 204, App
16	578	11	100.0	US-09-159-812-204	Sequence 204, App
17	578	11	100.0	US-09-636-215-204	Sequence 204, App
18	578	11	100.0	US-09-585-166A-204	Sequence 204, App
19	578	11	100.0	US-09-115-453-204	Sequence 204, App
20	578	11	100.0	US-09-688-489-204	Sequence 204, App
21	578	11	100.0	US-09-679-426-204	Sequence 204, App
22	578	11	100.0	US-09-759-143-204	Sequence 204, App
23	578	11	100.0	US-09-651-236-204	Sequence 204, App
24	578	11	100.0	US-09-949-016-18140	Sequence 18140, A
25	578	11	100.0	US-09-949-016-18400	Sequence 18400, A
26	578	11	100.0	US-09-949-016-21089	Sequence 21089, A
27	578	11	100.0	US-09-949-016-21090	Sequence 21090, A
28	578	11	100.0	US-09-949-016-26016	Sequence 26016, A
29	578	11	100.0	US-09-949-016-30150	Sequence 30150, A
30	578	11	100.0	US-09-949-016-33506	Sequence 33506, A
31	578	11	100.0	US-09-949-016-46876	Sequence 46876, A
32	578	11	100.0	US-09-949-016-50932	Sequence 50932, A
33	578	11	100.0	US-09-949-016-50933	Sequence 50933, A
34	578	11	100.0	US-09-949-016-50986	Sequence 50986, A
35	578	11	100.0	US-09-949-016-50987	Sequence 50987, A
36	578	11	100.0	US-09-949-016-53472	Sequence 53472, A
37	578	11	100.0	US-09-949-016-55344	Sequence 55344, A
38	578	11	100.0	US-09-949-016-55395	Sequence 55395, A
39	578	11	100.0	US-09-949-016-68211	Sequence 68211, A
40	578	11	100.0	US-09-949-016-68211	Sequence 68211, A
41	578	11	100.0	US-09-949-016-74618	Sequence 74618, A
42	578	11	100.0	US-09-949-016-80490	Sequence 80490, A
43	578	11	100.0	US-09-949-016-80491	Sequence 80491, A
44	578	11	100.0	US-09-949-016-80492	Sequence 80492, A
45	578	11	100.0	US-09-949-016-80493	Sequence 80493, A
46	578	11	100.0	US-09-949-016-87515	Sequence 87515, A
47	578	11	100.0	US-09-949-016-88349	Sequence 88349, A
48	578	11	100.0	US-09-949-016-88350	Sequence 88350, A
49	578	11	100.0	US-09-949-016-90108	Sequence 90108, A
50	578	11	100.0	US-09-949-016-91331	Sequence 91331, A
51	578	11	100.0	US-09-949-016-92251	Sequence 92251, A
52	578	11	100.0	US-09-949-016-107595	Sequence 107595, A
53	578	11	100.0	US-09-949-016-107596	Sequence 107596, A
54	578	11	100.0	US-09-949-016-112265	Sequence 112265, A
55	578	11	100.0	US-09-949-016-124285	Sequence 124285, A
56	578	11	100.0	US-09-949-016-124406	Sequence 124406, A
57	578	11	100.0	US-09-949-016-128965	Sequence 128965, A
58	578	11	100.0	US-09-949-016-13547	Sequence 13547, A
59	578	11	100.0	US-09-949-016-135316	Sequence 135316, A
60	578	11	100.0	US-09-949-016-135317	Sequence 135317, A
61	578	11	100.0	US-09-949-016-139202	Sequence 139202, A
62	578	11	100.0	US-09-949-016-142600	Sequence 142600, A
63	578	11	100.0	US-09-949-016-146476	Sequence 146476, A
64	578	11	100.0	US-09-949-016-146929	Sequence 146929, A
65	578	11	100.0	US-09-949-016-150808	Sequence 150808, A
66	578	11	100.0	US-09-949-016-150809	Sequence 150809, A
67	578	11	100.0	US-09-949-016-159317	Sequence 159317, A
68	578	11	100.0	US-09-949-016-160404	Sequence 160404, A
69	578	11	100.0	US-09-949-016-164192	Sequence 164192, A
70	578	11	100.0	US-09-949-016-165538	Sequence 165538, A
71	578	11	100.0	US-09-949-016-171890	Sequence 171890, A
72	578	11	100.0	US-09-949-016-172166	Sequence 172166, A

ALIGNMENTS

RESULT 1
US-09-513-999C-15732
; Sequence 15732, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 15732
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 152
; OTHER INFORMATION: y=c or t

US-09-513-999C-15732
Query Match 100.0%; Score 11; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||
Db 27 TAACTATATGA 37

RESULT 2
US-09-513-999C-32605
; Sequence 32605, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 32605
; LENGTH: 220
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 134
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 140
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 212

; OTHER INFORMATION: v=a or c or g
US-09-513-999C-32605

Query Match 100.0%; Score 11; DB 4; Length 220;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||
Db 201 TAACTATATGA 211

RESULT 3
US-09-513-999C-12608
; Sequence 12608, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 12608
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-12608

Query Match 100.0%; Score 11; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||
Db 196 TAACTATATGA 206

RESULT 4
US-09-270-767-31378
; Sequence 31378, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31378
; LENGTH: 284
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-31378

Query Match 100.0%; Score 11; DB 4; Length 284;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||
Db 259 TAACTATATGA 269

RESULT 5
US-09-328-352-1418


```
; Sequence 1418, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1418
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1418

Query Match      100.0%; Score 11; DB 4; Length 333;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      168 TAACTATATGA 178

RESULT 6
US-09-543-681A-3082
; Sequence 3082, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3082
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-3082

Query Match      100.0%; Score 11; DB 4; Length 336;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      237 TAACTATATGA 247

RESULT 7
US-09-543-681A-994
; Sequence 994, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 994
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-994
```

```
Query Match      100.0%; Score 11; DB 4; Length 348;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      6 TAACTATATGA 16

RESULT 8
US-09-134-000C-1499
; Sequence 1499, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1499
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1499
```

```
Query Match      100.0%; Score 11; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      156 TAACTATATGA 166
```

```
RESULT 9
US-09-313-294A-5027/c
; Sequence 5027, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5027
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700349462H1
; NAME/KEY: unsure
; LOCATION: 2, 58, 90, 116, 146, 163, 168, 183, 189, 205, 209, 212, 228, 232-233, 245,
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-5027
```

```
Query Match      100.0%; Score 11; DB 4; Length 397;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||||
Db      132 TAACTATATGA 122
```

```
RESULT 10
US-09-621-976-3695
; Sequence 3695, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Joubert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 3695
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CD3
; LOCATION: 132..284
US-09-621-976-3695

Query Match      100.0%; Score 11; DB 4; Length 422;
Best Local Similarity 100.0%; Pred. NO. 5.8e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      360 TAACTATATGA 370

RESULT 11
US-09-513-999C-11380/c
; Sequence 13380, Application US/09513999C
; Patent No. 6781961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 13380
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 252
; OTHER INFORMATION: w=a or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 274__feature
; OTHER INFORMATION: w=a or t
US-09-513-999C-11380

Query Match      100.0%; Score 11; DB 4; Length 460;
Best Local Similarity 100.0%; Pred. NO. 5.8e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      381 TAACTATATGA 371

RESULT 12
US-09-270-767-7858
; Sequence 7858, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7858
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-7858

Query Match      100.0%; Score 11; DB 4; Length 542;
Best Local Similarity 100.0%; Pred. NO. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      122 TAACTATATGA 132

RESULT 13
US-09-270-767-23140
; Sequence 23140, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23140
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-23140

Query Match      100.0%; Score 11; DB 4; Length 542;
Best Local Similarity 100.0%; Pred. NO. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      122 TAACTATATGA 132

RESULT 14
US-08-480-784-29/c
; Sequence 29, Application US/08480784
; Patent No. 5693473
; GENERAL INFORMATION:
; APPLICANT: Skolnick, Mark H.
; APPLICANT: Goldgar, David E.
; APPLICANT: Miki, Yoshio
; APPLICANT: Swenson, Jeff
; APPLICANT: Kamb, Alexander
; APPLICANT: Harshman, Keith D.
; APPLICANT: Shattuck-Bidens, Donna M.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Wiseman, Roger W.
; APPLICANT: Futreal, P. Andrew
; TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
; NUMBER OF SEQUENCES: 85
```

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
;; STREET: 1201 New York Avenue, N.W., Suite 1000
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20005
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/480,784
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/409,305
;; FILING DATE: 24-MAR-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/348,824
;; FILING DATE: 29-NOV-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/308,104
;; FILING DATE: 16-SEP-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/300,266
;; FILING DATE: 02-SEP-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/289,221
;; FILING DATE: 12-AUG-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ihnen, Jeffrey L.
;; REGISTRATION NUMBER: 28,957
;; REFERENCE/DOCKET NUMBER: 24884-109347
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-962-4810
;; TELEFAX: 202-962-8300
;; INFORMATION FOR SEQ ID NO: 29:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 561 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; US-08-480-784-29

Query Match 100.0%; Score 11; DB 1; Length 561;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 410 TAACTATATGA 400

RESULT 15
US-08-483-553-29/c
; Sequence 29, Application US/08483553
; Patent No. 5709999
; GENERAL INFORMATION:
; APPLICANT: Skolnick, Mark H.
; APPLICANT: Goldgar, David E.
; APPLICANT: Miki, Yoshio
; APPLICANT: Swenson, Jeff
; APPLICANT: Kamb, Alexander
; APPLICANT: Harshman, Keith D.
; APPLICANT: Shattuck-Bidens, Donna M.
; APPLICANT: Tavtigian, Sean V.

;; APPLICANT: Wiseman, Roger W.
;; APPLICANT: Futreal, P. Andrew
;; TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
;; TITLE OF INVENTION: Susceptibility Gene
;; NUMBER OF SEQUENCES: 85
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
;; STREET: 1201 New York Avenue, N.W., Suite 1000
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20005
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/483,553
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/409,305
;; FILING DATE: 24-MAR-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/348,824
;; FILING DATE: 29-NOV-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/308,104
;; FILING DATE: 16-SEP-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/300,266
;; FILING DATE: 02-SEP-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/289,221
;; FILING DATE: 12-AUG-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ihnen, Jeffrey L.
;; REGISTRATION NUMBER: 28,957
;; REFERENCE/DOCKET NUMBER: 24884-109347
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-962-4810
;; TELEFAX: 202-962-8300
;; INFORMATION FOR SEQ ID NO: 29:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 561 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; US-08-483-553-29

Query Match 100.0%; Score 11; DB 1; Length 561;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 410 TAACTATATGA 400

RESULT 16
US-08-487-002-29/c
; Sequence 29, Application US/08487002
; Patent No. 5710001
; GENERAL INFORMATION:
; APPLICANT: Shattuck-Bidens, Donna M.
; APPLICANT: Simard, Jacques
; APPLICANT: Eml, Mitsuru

;; APPLICANT: Nakamura, Yusuke
;; APPLICANT: Durocher, Francine
;; TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
;; TITLE OF INVENTION: Susceptibility Gene
;; NUMBER OF SEQUENCES: 85
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
;; STREET: 1201 New York Avenue, N.W., Suite 1000
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20005
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/487,002
;; FILING DATE:
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/409,305
;; FILING DATE: 24-MAR-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/348,824
;; FILING DATE: 29-NOV-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/308,104
;; FILING DATE: 16-SEP-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/300,266
;; FILING DATE: 02-SEP-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/289,221
;; FILING DATE: 12-AUG-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ihnen, Jeffrey L.
;; REGISTRATION NUMBER: 28,957
;; REFERENCE/DOCKET NUMBER: 24884-109347
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-962-4810
;; TELEFAX: 202-962-8300
;; INFORMATION FOR SEQ ID NO: 29:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 561 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; US-08-487-002-29

Query Match 100.0%; Score 11; DB 1; Length 561;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
Db 410 TAACTATATGA 400

RESULT 17
US-08-483-554B-29/c
; Sequence 29, Application US/08483554B
; Patent No. 5747282
; GENERAL INFORMATION:
; APPLICANT: Skolnick, Mark H.
; APPLICANT: Goldgar, David E.
; APPLICANT: Miki, Yoshio

;; APPLICANT: Swenson, Jeff
;; APPLICANT: Kamb, Alexander
;; APPLICANT: Harshman, Keith D.
;; APPLICANT: Shattuck-Eidens, Donna M.
;; APPLICANT: Tavtigian, Sean W.
;; APPLICANT: Wiseman, Roger W.
;; APPLICANT: Futreal, P. Andrew
;; TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
;; TITLE OF INVENTION: Susceptibility Gene
;; NUMBER OF SEQUENCES: 85
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
;; STREET: 1201 New York Avenue, N.W., Suite 1000
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20005
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/483,554B
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/409,305
;; FILING DATE: 24-MAR-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/348,824
;; FILING DATE: 29-NOV-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/308,104
;; FILING DATE: 16-SEP-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/300,266
;; FILING DATE: 02-SEP-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/289,221
;; FILING DATE: 12-AUG-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ihnen, Jeffrey L.
;; REGISTRATION NUMBER: 28,957
;; REFERENCE/DOCKET NUMBER: 24884-109347
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-962-4810
;; TELEFAX: 202-962-8300
;; INFORMATION FOR SEQ ID NO: 29:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 561 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; US-08-483-554B-29

Query Match 100.0%; Score 11; DB 1; Length 561;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGA 11
Db 410 TAACTATATGA 400

RESULT 18
US-08-488-011B-29/c
; Sequence 29, Application US/08488011B

Patent No. 5753441
GENERAL INFORMATION:
APPLICANT: Skolnick, Mark H.
APPLICANT: Goldgar, David E.
APPLICANT: Miki, Yoshio
APPLICANT: Swenson, Jeff
APPLICANT: Kamb, Alexander
APPLICANT: Harshman, Keith D.
APPLICANT: Shattuck-Eidens, Donna M.
APPLICANT: Tavtigian, Sean V.
APPLICANT: Wiseman, Roger W.
APPLICANT: Futreal, P. Andrew
TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
TITLE OF INVENTION: Susceptibility Gene
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,011B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409,305
FILING DATE: 24-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348,824
FILING DATE: 29-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/308,104
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,266
FILING DATE: 02-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/289,221
FILING DATE: 12-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24884-109347-09
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 561 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-488-011B-29

Query Match 100.0%; Score 11; DB 1; Length 561;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 410 TAACTATATGA 400

RESULT 19
US-08-850-727-29/c
Sequence 29, Application US/08850727
Patent No. 6162897
GENERAL INFORMATION:
APPLICANT: Skolnick, Mark H.
APPLICANT: Goldgar, David E.
APPLICANT: Miki, Yoshio
APPLICANT: Swenson, Jeff
APPLICANT: Kamb, Alexander
APPLICANT: Harshman, Keith D.
APPLICANT: Shattuck-Eidens, Donna M.
APPLICANT: Tavtigian, Sean V.
APPLICANT: Wiseman, Roger W.
APPLICANT: Futreal, P. Andrew
TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
TITLE OF INVENTION: Susceptibility Gene
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/850,727
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/483,554
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348,824
FILING DATE: 29-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/308,104
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,266
FILING DATE: 02-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/289,221
FILING DATE: 12-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24884-109347
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 561 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-850-727-29

Query Match 100.0%; Score 11; DB 3; Length 561;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 410 TAACTATATGA 400

Query Match 100.0%; Score 11; DB 5; Length 561;
Best Local Similarity 100.0%; Freq. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 410 TAACTATATGA 400

RESULT 20

PCT-US95-10202-29/c
; Sequence 29, Application PC/TUS9510202
; GENERAL INFORMATION:
; APPLICANT: Shattuck-Bidens, Donna M.
; APPLICANT: Simard, Jacques
; APPLICANT: Emi, Mitsuru
; APPLICANT: Nakamura, Yuseke
; APPLICANT: Durocher, Francine
; TITLE OF INVENTION: In Vivo Mutations and Polymorphisms
; TITLE OF INVENTION: In the 17q-Linked Breast and Ovarian Cancer
; TITLE OF INVENTION: Susceptibility Gene
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10202
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07-JUN-1995
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/409,305
FILING DATE: 24-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348,824
FILING DATE: 29-NOV-1994
APPLICATION DATA:
APPLICATION NUMBER: US 08-308,104
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,266
FILING DATE: 02-SEP-1994
APPLICATION DATA:
APPLICATION NUMBER: US 08/289,221
FILING DATE: 12-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24884-109347
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 561 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
PCT-US95-10202-29

RESULT 21

PCT-US95-10203-29/c
; Sequence 29, Application PC/TUS9510203
; GENERAL INFORMATION:
; APPLICANT: Skolnick, Mark H.
; APPLICANT: Goldgar, David E.
; APPLICANT: Miki, Yoshio
; APPLICANT: Swenson, Jeff
; APPLICANT: Kamb, Alexander
; APPLICANT: Harshman, Keith D.
; APPLICANT: Shattuck-Bidens, Donna M.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Wiseman, Roger W.
; APPLICANT: Futreal, P. Andrew
; TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
; TITLE OF INVENTION: Susceptibility Gene
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10203
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07-JUN-1995
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/409,305
FILING DATE: 24-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348,824
FILING DATE: 29-NOV-1994
APPLICATION DATA:
APPLICATION NUMBER: US 08-308,104
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,266
FILING DATE: 02-SEP-1994
APPLICATION DATA:
APPLICATION NUMBER: US 08/289,221
FILING DATE: 12-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24884-109347
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 561 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
PCT-US95-10203-29

Query Match          100.0%; Score 11; DB 5; Length 561;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      410 TAACTATATGA 400

RESULT 22
PCT-US95-10220-29/c
; Sequence 29, Application PC/TUS9510220
; GENERAL INFORMATION:
; APPLICANT: Skolnick, Mark H.
; APPLICANT: Goldgar, David E.
; APPLICANT: Miki, Yoshio
; APPLICANT: Swenson, Jeff
; APPLICANT: Kamb, Alexander
; APPLICANT: Harshman, Keith D.
; APPLICANT: Shattuck-Eidens, Donna M.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Wiseman, Roger W.
; APPLICANT: Futreal, P. Andrew
; TITLE OF INVENTION: Method for diagnosing a
; TITLE OF INVENTION: Predisposition for Breast and Ovarian Cancer
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10220
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/409,305
; FILING DATE: 24-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,824
; FILING DATE: 29-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08-308,104
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300,266
; FILING DATE: 02-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/289,221
; FILING DATE: 12-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24884-109347
; TELECOMMUNICATION INFORMATION:
```

```
;
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 561 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
PCT-US95-10220-29

Query Match          100.0%; Score 11; DB 5; Length 561;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      410 TAACTATATGA 400

RESULT 23
US-09-328-352-3569/c
; Sequence 3569, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3569
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3569

Query Match          100.0%; Score 11; DB 4; Length 570;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      62 TAACTATATGA 52

RESULT 24
US-09-020-956-104/c
; Sequence 104, Application US/09020956
; Patent No. 6261562
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillin, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/020,956
; FILING DATE: 09-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C2
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 578 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
US-09-020-956-104

Query Match 100.0%; Score 11; DB 3; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 215 TAACTATATGA 205

RESULT 25
US-09-030-607-104/c
; Sequence 104, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 4300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C3
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 578 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
US-09-030-607-104
```

```
Query Match 100.0%; Score 11; DB 3; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 215 TAACTATATGA 205

RESULT 26
US-09-439-313-104/c
; Sequence 104, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapien
;
US-09-439-313-104

Query Match 100.0%; Score 11; DB 3; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 215 TAACTATATGA 205

RESULT 27
US-09-352-616A-104/c
; Sequence 104, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Yuqui
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapien
;
US-09-352-616A-104

Query Match 100.0%; Score 11; DB 3; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```

Qy      1 TAACTATATGA 11
Db      215 TAACTATATGA 205

RESULT 28
US-09-232-149A-104/c
; Sequence 104, Application US/09232149A
; Patent No. 6465611
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C6
; CURRENT APPLICATION NUMBER: US/09/232,149A
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-232-149A-104

Query Match      100.0%; Score 11; DB 3; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      215 TAACTATATGA 205

RESULT 29
US-09-159-812-104/c
; Sequence 104, Application US/09159812A
; Patent No. 6613872
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF
; TITLE OF INVENTION: PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.428C5
; CURRENT APPLICATION NUMBER: US/09/159,812A
; CURRENT FILING DATE: 1998-09-23
; NUMBER OF SEQ ID NOS: 306
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-159-812-104

Query Match      100.0%; Score 11; DB 4; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      215 TAACTATATGA 205

RESULT 30
US-09-636-215-104/c
; Sequence 104, Application US/09636215
; Patent No. 6620322
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.

Query Match      100.0%; Score 11; DB 4; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      215 TAACTATATGA 205

RESULT 31
US-09-685-166A-104/c
; Sequence 104, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-685-166A-104

Query Match      100.0%; Score 11; DB 4; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      215 TAACTATATGA 205

US-09-636-215-104
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-636-215-104

Query Match      100.0%; Score 11; DB 4; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      215 TAACTATATGA 205

```

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
| | | | | | | | | |
Db 215 TAACTATATGA 205

RESULT 32

US-09-115-453-104/c
; Sequence 104, Application US/09115453B
; Patent No. 6657056
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND
; TITLE OF INVENTION: METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C4
; CURRENT APPLICATION NUMBER: US/09/115,453B
; CURRENT FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-115-453-104

Query Match 100.0%; Score 11; DB 4; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
| | | | | | | | | |
Db 215 TAACTATATGA 205

RESULT 33

US-09-688-489-104/c
; Sequence 104, Application US/09688489
; Patent No. 6664377
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427D2
; CURRENT APPLICATION NUMBER: US/09/688,489
; CURRENT FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-688-489-104

Query Match 100.0%; Score 11; DB 4; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
| | | | | | | | | |
Db 215 TAACTATATGA 205

RESULT 34

US-09-679-426-104/c
; Sequence 104, Application US/09679426
; Patent No. 6759515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C20
; CURRENT APPLICATION NUMBER: US/09/679,426
; CURRENT FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 895
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-679-426-104

Query Match 100.0%; Score 11; DB 4; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
| | | | | | | | | |
Db 215 TAACTATATGA 205

RESULT 35

US-09-759-143-104/c
; Sequence 104, Application US/09759143
; Patent No. 6800746
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-759-143-104

Query Match 100.0%; Score 11; DB 4; Length 578;

```
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 215 TAACTATATGA 205

RESULT 36
US-09-651-236-104/c
; Sequence 104, Application US/09651236
; Patent No. 6818751
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42718C18
; CURRENT APPLICATION NUMBER: US/09/651,236
; CURRENT FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 865
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-651-236-104

Query Match 100.0%; Score 11; DB 4; Length 578;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 215 TAACTATATGA 205

RESULT 37
US-09-030-607-204/c
; Sequence 204, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 204:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-030-607-204

Query Match 100.0%; Score 11; DB 3; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 218 TAACTATATGA 208

RESULT 38
US-09-439-313-204/c
; Sequence 204, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 204
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(589)
; OTHER INFORMATION: n = A,T,C or G
US-09-439-313-204

Query Match 100.0%; Score 11; DB 3; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
Db 218 TAACTATATGA 208

RESULT 39
US-09-352-616A-204/c
; Sequence 204, Application US/09352616A
```

; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Yuqui
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.42708
; CURRENT APPLICATION NUMBER: US/09/352.616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 204
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(589)
; OTHER INFORMATION: n = A,T,C or G
US-09-352-616A-204

Query Match 100.0%; Score 11; DB 3; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
DB 218 TAACTATATGCA 208
|||||

RESULT 40

US-09-232-149A-204/c
; Sequence 204, Application US/09232149A
; Patent No. 6465611
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.42708
; CURRENT APPLICATION NUMBER: US/09/232.149A
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 204
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(589)
; OTHER INFORMATION: n = A,T,C or G
US-09-232-149A-204

Query Match 100.0%; Score 11; DB 3; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
DB 218 TAACTATATGCA 208
|||||

RESULT 41

US-09-159-812-204/c
; Sequence 204, Application US/09159812A
; Patent No. 6613872
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF
; TITLE OF INVENTION: PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.42805
; CURRENT APPLICATION NUMBER: US/09/159.812A
; CURRENT FILING DATE: 1998-09-23
; NUMBER OF SEQ ID NOS: 306
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 204
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(589)
; OTHER INFORMATION: n = A,T,C or G
US-09-159-812-204

Query Match 100.0%; Score 11; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
DB 218 TAACTATATGCA 208
|||||

RESULT 42

US-09-636-215-204/c
; Sequence 204, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Wang, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636.215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 204
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(589)
; OTHER INFORMATION: n = A,T,C or G
US-09-636-215-204

Query Match 100.0%; Score 11; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACTATATGCA 11
DB 218 TAACTATATGCA 208
|||||

```
RESULT 43
US-09-685-166A-204/c
; Sequence 204, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 204
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(589)
; OTHER INFORMATION: n = A,T,C or G
US-09-685-166A-204

Query Match      100.0%; Score 11; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||
Db      218 TAACTATATGA 208

RESULT 44
US-09-115-453-204/c
; Sequence 204, Application US/09115453B
; Patent No. 6657056
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND
; FILE REFERENCE: 210121.427C4
; CURRENT APPLICATION NUMBER: US/09/115,453B
; CURRENT FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 204
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(589)
; OTHER INFORMATION: n = A,T,C or G
US-09-115-453-204

Query Match      100.0%; Score 11; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||
Db      218 TAACTATATGA 208

RESULT 45
US-09-688-489-204/c
; Sequence 204, Application US/09688489
; Patent No. 6664377
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; FILE REFERENCE: 210121.427D2
; CURRENT APPLICATION NUMBER: US/09/688,489
; CURRENT FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 204
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(589)
; OTHER INFORMATION: n = A,T,C or G
US-09-688-489-204

Query Match      100.0%; Score 11; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
      |||
Db      218 TAACTATATGA 208

RESULT 46
US-09-679-426-204/c
; Sequence 204, Application US/09679426
; Patent No. 6759515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C20
; CURRENT APPLICATION NUMBER: US/09/679,426
; CURRENT FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 895
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 204
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(589)
; OTHER INFORMATION: n = A,T,C or G
US-09-679-426-204

Query Match      100.0%; Score 11; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGCA 11
Db      218 TAACTATATGCA 208

RESULT 47
US-09-759-143-204/c
; Sequence 204, Application US/09759143
; Patent No. 6800746
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 204
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(589)
; OTHER INFORMATION: n = A,T,C or G
US-09-759-143-204

Query Match      100.0%; Score 11; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGCA 11
Db      218 TAACTATATGCA 208

RESULT 48
US-09-651-236-204/c
; Sequence 204, Application US/09651236
; Patent No. 6818751
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
```

```
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.42718C18
; CURRENT APPLICATION NUMBER: US/09/651,236
; CURRENT FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 865
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 204
; LENGTH: 589
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(589)
; OTHER INFORMATION: n = A,T,C or G
US-09-651-236-204

Query Match      100.0%; Score 11; DB 4; Length 589;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGCA 11
Db      218 TAACTATATGCA 208

RESULT 49
US-09-949-016-18140/c
; Sequence 18140, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18140
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-18140

Query Match      100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGCA 11
Db      218 TAACTATATGCA 208
```

Db 206 TAACTATATGA 196

RESULT 50

US-09-949-016-18400/c
; Sequence 18400, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18400
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-18400

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

Db 365 TAACTATATGA 355

RESULT 51

US-09-949-016-21089
; Sequence 21089, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21089
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-21089

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

Db 132 TAACTATATGA 142

RESULT 52

US-09-949-016-21090

; Sequence 21090, Application US/09949016

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21090
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-21090

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

Db 238 TAACTATATGA 248

RESULT 53

US-09-949-016-26016
; Sequence 26016, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26016
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-26016

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11

Db 59 TAACTATATGA 69

RESULT 54

US-09-949-016-30150/c
; Sequence 30150, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30150
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-30150

Query Match          100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TAACTATATGCA 11
Db      232 TAACTATATGCA 222

RESULT 55
US-09-949-016-33506
; Sequence 33506, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33506
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-33506

Query Match          100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TAACTATATGCA 11
Db      12 TAACTATATGCA 22

RESULT 56
US-09-949-016-40797
; Sequence 40797, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40797
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-40797

Query Match          100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TAACTATATGCA 11
Db      400 TAACTATATGCA 410

RESULT 57
US-09-949-016-46876
; Sequence 46876, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46876
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-46876

Query Match          100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TAACTATATGCA 11
Db      533 TAACTATATGCA 543

RESULT 58
US-09-949-016-50932/c
; Sequence 50932, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
```


; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50932
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-50932

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 133 TAACTATATGA 123

RESULT 59

US-09-949-016-50933/c
; Sequence 50933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50933
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-50933

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 101 TAACTATATGA 91

RESULT 60

US-09-949-016-50986/c
; Sequence 50986, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50986
; LENGTH: 601
; TYPE: DNA

; ORGANISM: Human
; US-09-949-016-50986

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 133 TAACTATATGA 123

RESULT 61

US-09-949-016-50987/c
; Sequence 50987, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50987
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-50987

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||||
Db 101 TAACTATATGA 91

RESULT 62

US-09-949-016-53472/c
; Sequence 53472, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53472
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-53472

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;

```
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
   |||||
Db 301 TAACTATATGA 291

RESULT 63
US-09-949-016-55344
; Sequence 55344, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55344
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-55344

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
   |||||
Db 87 TAACTATATGA 97

RESULT 64
US-09-949-016-55395/C
; Sequence 55395, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55395
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-55395

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
   |||||
Db 383 TAACTATATGA 373

RESULT 65
US-09-949-016-60118
; Sequence 60118, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60118
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-60118

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
   |||||
Db 85 TAACTATATGA 95

RESULT 66
US-09-949-016-68211
; Sequence 68211, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68211
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-68211

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
   |||||
Db 481 TAACTATATGA 491

RESULT 67
US-09-949-016-74618
; Sequence 74618, Application US/09949016
```

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74618
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-74618

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
|||||
Db 59 TAACTATATGCA 69

RESULT 68
US-09-949-016-80490/c
; Sequence 80490, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80490
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-80490

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
|||||
Db 593 TAACTATATGCA 593

RESULT 69
US-09-949-016-80491/c
; Sequence 80491, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80491
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-80491

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
|||||
Db 581 TAACTATATGCA 571

RESULT 70

US-09-949-016-80492/c
; Sequence 80492, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80492
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-80492

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGCA 11
|||||
Db 431 TAACTATATGCA 421

RESULT 71

US-09-949-016-80493/c
; Sequence 80493, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80493
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-88349

Query Match          100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      426 TAACTATATGA 416

RESULT 72
US-09-949-016-87515
; Sequence 87515, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87515
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-87515

Query Match          100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      483 TAACTATATGA 493

RESULT 73
US-09-949-016-88349/c
; Sequence 88349, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88349
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-88349

Query Match          100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      270 TAACTATATGA 260

RESULT 74
US-09-949-016-88350/c
; Sequence 88350, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88350
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-88350

Query Match          100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TAACTATATGA 11
Db      270 TAACTATATGA 260

RESULT 75
US-09-949-016-90108
; Sequence 90108, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 90108
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-90108
```

US-09-949-016-90108

Query Match 100.0%; Score 11; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAACTATATGA 11
|||
Db 117 TAACTATATGA 127

Search completed: August 30, 2005, 13:00:03
Job time : 2.93262 secs

This Page Blank (uspto)